

Matches 173; Conservative 0; Mismatches 176; Indels 3; Gaps 1;
 QY 63 GCGCCCGGACGATGACCAAGAGGAGTGGCCAGCGCATGACAGAGGACCGGCGCCCT 122
 Db 565 GGGCGGTGTGTGGATCATAGAGTCCGGCCGAACTGACATGACGCGCCGACGCT 624
 QY 123 CGGCGGTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 182
 Db 625 GTGCTGAAATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684
 QY 183 CGAGGTGCTGGGCTGACCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 242
 Db 685 CGAGGAGATGCGCGCTGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
 QY 243 CACCCCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 302
 Db 742 GCGCGGTGAGGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
 QY 303 CCGGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
 Db 802 CGAATACGTGACGAGTGTGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
 QY 363 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 414
 Db 862 CGAGGCGAAGGTGCGGTGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913

RESULT 8

US-09-252-991A-3645
 ; Sequence 3645, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 3645
 ; LENGTH: 939
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3645

Query Match 12.2%; Score 52; DB 4; Length 939;
 Best Local Similarity 47.3%; Pred. No. 0.045;

Matches 194; Conservative 0; Mismatches 210; Indels 6; Gaps 1;

QY 14 CGGGGCTGGGGCGGACATGAGGCGGATGCTCTCAACCGCGGAGCGGCGGAGC 73
 Db 348 CGTCCCGGAGCGCCGCTGTGAGCGGCTTACGCGGAGCATGACCGCTTTGTGCGCA 407
 QY 74 TGACCAAGAGGAGTGGCGGAGGCGGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 133
 Db 408 TCCTCCAGGCGGAGGCGGAGGCGGCTGCAACTGCTCAGACGAGTGAAGGCGCTGTGCC 467
 QY 134 AGAGCGGAGAGGAGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 193
 Db 468 TGTTCCTAGGCTGT 527
 QY 194 GCTTCGACCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 253
 Db 528 GCGCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 587
 QY 254 GAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313
 Db 588 GGTGACCTTACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 647

QY 314 ACAGGACATGAAAGCGGCGCATATGCGCCCTTATCTGTGAGGCGCGGTGAGCGGAGCAAG 373
 Db 648 CCGGCGAGCTGAAGG-----CATTTAGGCGGAGCTGTAAGAGCGCGGTGAGGAGCAAGA 701
 QY 374 CCGGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 423
 Db 702 CCGGCGCCTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751

RESULT 9

US-09-252-991A-3697
 ; Sequence 3697, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 3697
 ; LENGTH: 1881
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3697

Query Match 12.2%; Score 52; DB 4; Length 1881;
 Best Local Similarity 47.3%; Pred. No. 0.047;
 Matches 194; Conservative 0; Mismatches 210; Indels 6; Gaps 1;

QY 14 CGGGGCTGGGGCGGACATGAGGCGGATGCTCTCAACCGCGGAGCGGCGGAGC 73
 Db 419 CGTCCCGGAGCGCCGCTGTGAGCGGCTTACGCGGAGCATGACCGCTTTGTGCGCA 478
 QY 74 TGACCAAGAGGAGTGGCGGAGGCGGATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 133
 Db 479 TCCTCCAGGCGGAGGCGGAGGCGGCTGCAACTGCTCAGACGAGTGAAGGCGCTGTGCC 538
 QY 134 AGAGCGGAGAGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 193
 Db 539 TGTTCCTAGGCTGT 598
 QY 194 GCTTCGACCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 253
 Db 599 GCGCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658
 QY 254 GAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 313
 Db 659 GGTGACCTTACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718
 QY 314 ACAGGAGATGAAAGCGGCGCATATGCGCCCTTATCTGTGAGGCGCGGTGAGGAGGAGCAAG 373
 Db 719 CCGGCGAGCTGAAG-----CATTTAGGCGGAGCTGTAAGAGCGCGGTGAGGAGCAAGA 772
 QY 374 CCGGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 423
 Db 773 CCGGCGCCTCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 822

RESULT 10

US-09-252-991A-3812/c
 ; Sequence 3812, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3812
LENGTH: 1998
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3812

Query Match 12.2%; Score 52; DB 4; Length 1998;
Best Local Similarity 47.3%; Pred. No. 0.047;
Matches 194; Conservative 0; Mismatches 210; Indels 6; Gaps 1;

14 CGGGGCTGGGGCGGGGACATGAGCGGCGCATACGCTCTACCGCGCGGAGCGCGGAC 73
1568 CGCTGCCCGGAGCGCGCTGTGAGCGCTACGCTGCGGACATGACCGTTTGTCTGCGCA 1509
74 TGACCAAGAGCGAGTTTGCCAGCGGCGATCCAGAGACCGGCGCACCGTGGCGGTGG 133
1508 TCCTCCAGGCGGAGCGGCGGCTGCGCAACTGCTCAGACGCTGCGAGGCGCTGTGCC 1449
134 AGGACGGCGAAGAACCGGCGCGGACGCGGACCTGCTTGGCGCGGTGGCGGAGTGTGG 193
1448 TGTTCCTCAGCTGT 1389
194 GCCTGACCTTGACGAGAACCTTCCGCGCGCGGCGGAGCTGTGCGCGCGCGCGCG 253
1388 GCCCGCTGCGGAGTCACTACGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1329
254 CGACCCCAACATGAGACCTGAGCGGAGAAATGAGCTGTGTGCGACCGACCCCAAGCTGG 313
1328 GGGTCACTTACAGCGCGGCGGAGCGGAGCTGTGCGGAGCGGCGGAGCGGCGGAGTGG 1269
314 ACGAGGACATGAGAGCGGCGGATCGGCGCTTATCTGTGAGGCGCGGTGAGCGGCGGAG 373
1268 CGGCGAGCTGAGAG-----CATCTACGCGGAGCGGCGGAGAGCGGCGGAGCGGAG 1215
374 CGGCGGCGATCGAGGAGAACCAAGCGGCGTCTATGAGCTGTGTGCGCGCGGAGC 423
1214 CCGCGCGCTCTGCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1165

RESULT 11
US-08-804-227C-7
Sequence 7, Application US/08804227C

GENERAL INFORMATION:
Patent No. 5876991
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhnstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sulton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII (DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4437 base pairs
TYPE: nucleic acid
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002

FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036

FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284

FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071

FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830

US-08-804-227C-7

Query Match 12.0%; Score 51; DB 2; Length 4437;
Best Local Similarity 47.9%; Pred. No. 0.094;
Matches 147; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

14 CGGGGCTGGGGCGGGGACATGAGCGGCGCATACGCTCTACCGCGCGGAGCGCGGAC 1165
2549 CCGAGGCG 1165
74 TGACCAAGAGCGAGTTTGCCAGCGGCGGATCCAGAGAGCGGCGCACCGTGTGCGCGGT 1165
2609 TGTATGAGAGAGCG 1165
134 AGGACGGCGAAGAACCGGCGCGGACGAGCGGAGCTGTGTGCGCGCGCGCGCGCGCG 1165
2609 CCGAGGCGGAGCTGT 1165
194 GCCTGACCTTGACGAGAACCTTCCGCGCGCGGCGGAGCTGTGCGCGCGCGCGCGCG 1165
26129 CCGCGGACAAAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1165
254 CGACCCCAACATGAGACCTGAGCGGAGAAATGAGCTGTGTGCGACCGGCGGAGCG 1165
26189 CCCACCCGAGCTGACCGCGCGGAGCGGCGGAGCTGTGCGGCTACCTTATCACC 1165
314 ACGAGGA 320
26249 TCGAGCA 26255

RESULT 12
US-08-804-198-1
Sequence 1, Application US/08804198

GENERAL INFORMATION:
Patent No. 5945320
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhnstoss, Stuart A.
APPLICANT: Rao, Nagaraia R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138

STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-198-1

Query Match 12.0%; Score 51; DB 2; Length 44377;
Best Local Similarity 47.9%; Pred. No. 0.094;
Matches 147; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

14 CCGGCTGAGGCGCGGCACATGAGCGCATACCTCTCAACCGCGCGAGCGCGCGGAC 73
25949 CCGAGGGGCGCGCGCGCGGCGATCTCTCTGCGCATACGCGACCAACGCGCAC 26008
74 TGACCAAGAGCGAGTGGCGCAGCGCATCCAGAGAGCGCGCGCACCGTGGCGCGG 133
26009 TCGTATCGAGAGCGCGCGAGCGCACCGCGCGCGGAACTGCTCGCGCGCGAGCG 26068
134 AGGACGGCAAGAACCGCGCGCGAGCGGACCTGTGCGCGCGTGGCGCGAGTCTCG 193
26069 CCGAGCGCGACCTGTGCTCGAGAGTGTGCGACAGAGTGAACGTCGCGCGATGAT 26128
194 GCTTCGACCTGACGAGCGCGCGCGCGCGCGAGTGTGCGCGCGCGCGCGTCAAC 253
26129 CCGGCGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26188
254 GCACCCCAACCATGAGCTGAGCGAGAGAAATGAGCTGTGCGCGCGCGCGCGCGCG 313
26189 CCCACCGCGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 26248
314 ACAGGGA 320
26249 TCGAGCA 26255

RESULT 13
US-09-266-965-22
Sequence 22; Application US/09266965
Patent No. 6495348
GENERAL INFORMATION:
APPLICANT: Sherman, D
APPLICANT: Mao, Y
APPLICANT: Varoglu, M
APPLICANT: He, M
APPLICANT: Sheldon, P
TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
FILE REFERENCE: 600,456US1
CURRENT APPLICATION NUMBER: US/09/266,965
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: US 08/624,447
EARLIER FILING DATE: 1996-08-19
EARLIER APPLICATION NUMBER: PCT/US94/11279
EARLIER FILING DATE: 1994-10-06
EARLIER APPLICATION NUMBER: US 08/133,963
EARLIER FILING DATE: 1993-10-07
NUMBER OF SEQ ID NOS: 145
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22
LENGTH: 1224
TYPE: DNA
ORGANISM: Streptomyces lavendulae
US-09-266-965-22

Query Match 11.8%; Score 50.4; DB 4; Length 1224;
Best Local Similarity 47.5%; Pred. No. 0.097;
Matches 150; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

99 CATCCAGAGAGCG 158
861 CCGTCCAGAGAGCG 920
159 CCGGACCTGCTGCG 218
921 GAAGACCGGAGAGTCCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 980
219 CCGCGAGGCTGCG 278
981 CCGCGAGGCTGCG 1040
279 GGAATCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 338
1041 CCGCATTCACCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1100
339 CCGCTTAATCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 398
1101 CCGCTTCTGACGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1160
399 GCTCATGACCTGTTG 414
1161 GCGGAGGACGACGATC 1176

RESULT 14
US-09-902-540-8618
Sequence 8618; Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825

THIS PAGE BLANK (USPTO)

Mon Apr 18 09:47:30 2005

us-09-855-340b-2.rnpb

Pa

GenCore version 5.1.6
Copyright (c) 1993 - 2005. CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 19:01:23 ; Search time 315.016 Seconds
(without alignments)
8204.076 Million cell updates/sec

Title: US-09-855-340b-2

Perfect score: 426
Sequence: 1 atcgcaacacacacccggggct.....accgttcgcccggagctga 426

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10F_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	426	100.0	426	9	US-09-855-340-2
2	64.4	15.1	1056	10	US-09-953-348-66
3	64.4	15.1	1056	15	US-10-267-255-66
4	64.4	15.1	53500	10	US-09-953-348-76
5	64.4	15.1	53500	15	US-10-267-255-76
6	63.4	14.9	1033	18	US-10-437-963-54677
7	60.6	14.2	683	17	US-10-425-115-73639
8	60.6	14.2	2129	17	US-10-425-114-25190
9	58.6	13.8	1284	15	US-10-156-761-3520
10	58.6	13.8	1365	15	US-10-156-761-43348
11	58.6	13.8	9025608	15	US-10-156-761-1

12	58.6	13.8	9025608	15	US-10-156-761-1
13	58	13.6	1332	15	US-10-156-761-4084
14	57.8	13.6	1108	18	US-10-437-963-75092
15	57.8	13.6	1659	15	US-10-156-761-6116
16	57.6	13.5	1882	18	US-10-767-701-12741
17	57	13.4	15738	16	US-10-329-079-46
18	57	13.4	61944	16	US-10-329-079-34
19	56.8	13.3	972	17	US-10-107-431-246
20	56.8	13.3	45055	17	US-10-107-431-277
21	56.4	13.2	1262	17	US-10-425-114-3009
22	56.4	13.2	1411	18	US-10-425-115-102021
23	56.2	13.2	966	18	US-10-437-963-62311
24	56.2	13.2	3018	18	US-10-437-963-8241
25	56	13.1	662	18	US-10-653-047-5009
26	55.6	13.1	1620	17	US-10-369-493-31859
27	55.4	13.0	88421	9	US-09-976-059-1
28	54.8	12.9	546	18	US-10-437-963-22108
29	54.8	12.9	137560	18	US-10-481-112-1
30	54.6	12.8	915	15	US-10-156-761-4240
31	54.6	12.8	1509	15	US-10-156-761-6656
32	54.4	12.8	2797	18	US-10-437-963-80187
33	54.4	12.8	2706	18	US-10-437-963-80187
34	54.2	12.7	1234	17	US-10-425-115-26760
35	54.2	12.7	1230	18	US-10-425-114-14401
36	54.2	12.7	1239	18	US-10-425-115-177618
37	54.2	12.7	36321	15	US-10-187-267A-1
38	54	12.7	5127	15	US-10-156-761-3630
39	53.8	12.6	861	16	US-10-190-305A-16
40	53.8	12.6	866	16	US-10-241-009-23
41	53.8	12.6	866	16	US-10-190-4348-23
42	53.8	12.6	975	18	US-10-437-963-61696
43	53.8	12.6	1380	15	US-10-156-761-5617
44	53.8	12.6	1953	18	US-10-437-963-54293
45	53.8	12.6	1977	17	US-10-369-493-35601

ALIGNMENTS

RESULT 1
US-09-855-340-2
Sequence 2, Application US/09855340
Patent No. US20020076788A1
GENERAL INFORMATION:
APPLICANT: Hosted, Jr., Thomas J.
TITLE OF INVENTION: Isolation of Micromonospora carbonacea var affinis
TITLE OF INVENTION: PMMP integrase and use of integrating function
TITLE OF INVENTION: site-specific integration into Micromonospora
TITLE OF INVENTION: halophila and Micromonospora carbonacea chrom
FILE REFERENCE: IN01164K
CURRENT APPLICATION NUMBER: US/09/855,340
CURRENT FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/204,670
PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 426
TYPE: DNA
ORGANISM: Micromonospora carbonacea
US-09-855-340-2

Query Match
Best Local Similarity 100.0%; Pred. No. 1.3e-97;
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0

DB 1
1 ATGCGCAACACACCGGGGCTGGGGCGGCGCATGCGCGCATCTCTCCACCGCC
61 GAGCGCGCGGAGTGAACGAAGCGATTGGCCGCGCATCCGAAGAGCGGGGCC

Db 61 GAGCGCGCCGAGTACCAAGAGGAGTGTGGCCAGCGGCATCCAGAAAGACCGGGCCACC 120
 QY 121 GTGCGCGGTGGAGAGACCGGCAAGACCGGCGGAGCGGAGCCTGTTGGCCCGGTC 180
 Db 121 GTGCGCGGTGGAGAGACCGGCAAGACCGGCGGAGCGGAGCCTGTTGGCCCGGTC 180
 QY 181 GCCCAGGTGCTGGGCTCGACCTCGAAGAGCCCTGGCCCGGAGTCTGGCCCGGCG 240
 Db 181 GCCCAGGTGCTGGGCTCGACCTCGAAGAGCCCTGGCCCGGAGTCTGGCCCGGCG 240
 QY 241 GTGACCCCGGCGAGCGACCCCAACCATGAGCCTGAGAGAGAAATCGAGCTGGTCCGAC 300
 Db 241 GTGACCCCGGCGAGCGACCCCAACCATGAGCCTGAGAGAGAAATCGAGCTGGTCCGAC 300
 QY 301 GACCCCAAGCTGAGCGAGGACATGAAGCGCGCATGCGCCTTAATCTTGAAGCGCGCT 360
 Db 301 GACCCCAAGCTGAGCGAGGACATGAAGCGCGCATGCGCCTTAATCTTGAAGCGCGCT 360
 QY 361 GAGCGCGAGAAAGCGCGCGGCTGAGAGAAACCAAGCGGCTCATGACCTGTTCCGCGG 420
 Db 361 GAGCGCGAGAAAGCGCGCGGCTGAGAGAAACCAAGCGGCTCATGACCTGTTCCGCGG 420
 QY 421 AGCTGA 426
 Db 421 AGCTGA 426

RESULT 2
 US-09-953-348-66
 ; Sequence 66, Application US/09953348
 ; Publication No. US20030134398A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, David. H
 ; APPLICANT: Mao, Yingqing
 ; APPLICANT: Varoglu, Mustafa
 ; APPLICANT: He, Min
 ; APPLICANT: Sheldon, Paul
 ; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
 ; FILE REFERENCE: 600.530US1
 ; CURRENT APPLICATION NUMBER: US/09/953.348
 ; PRIOR FILING DATE: 2001-09-12
 ; PRIOR APPLICATION NUMBER: PCT/US00/06394
 ; PRIOR FILING DATE: 2000-03-10
 ; PRIOR APPLICATION NUMBER: 09/266965
 ; NUMBER OF SEQ ID NOS: 153
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 66
 ; LENGTH: 1056
 ; TYPE: DNA
 ; ORGANISM: Streptomyces lavendulae
 ; US-09-953-348-66

Query Match 15.1%; Score 64.4; DB 10; Length 1056;
 Best Local Similarity 51.0%; Pred. No. 5.5e-07;
 Matches 152; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 92 CCAGGCGCATTCAGAGAGACCGGCGCACCGTGGCGCGGTGGAGAGACGGCAAGACCGGC 151
 Db 683 CCGGCGCGCGGCTCGCGGACCGGTGGAATCTGCGCGGCGCATCTTTCGAGAACCATCC 742
 QY 152 CCGAGCGAGCGGACCTGCTTGGCCCGGCTGCGCCAGGAGTGGCGGCTCGACCTCGAGAG 211
 Db 743 CCGAGCGGCGCGGACGCTTACCTATTAAGACAGTGTCTGACGAGCTGGAGCGAGAGAG 802
 QY 212 CCTGCGCGCGGAGGTCTGCGCGCGGCGGATCAACCGCGAGCGACCCCAACCATGAGC 271
 Db 803 TCGTAGCGATCTCTCGCGCGGATGCGACCGCATGAAGCGGAGCTCCGGCTCTGGTGA 862
 QY 272 TGGAGCGAGAAATCGAGCTGTGTCGCGACCGACCCCAAGTGAAGAGAGCATGAAGCGGC 331
 Db 863 TCGACAACTCTATCGAGCGAGCGCGCGCGCATGAGCGCTCTTTCGAGACCTGCTGGC 922

QY 332 GCATATCGCCCTTAATCTGAGCGCGCGGAGCGGAGCAAGAGGCGCGGATGAGGA 389
 Db 923 TCGTCTCGTGGCGCGCGCGGAGCGCTCGAGAGCGAATTCGCGCGCTGCTGAGAA 980

RESULT 3
 US-10-267-255-66
 ; Sequence 66, Application US/10267255
 ; Publication No. US20030124689A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D
 ; APPLICANT: Mao, Y
 ; APPLICANT: Varoglu, M
 ; APPLICANT: He, M
 ; APPLICANT: Sheldon, P
 ; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
 ; FILE REFERENCE: 600.456US1
 ; CURRENT APPLICATION NUMBER: US/10/267.255
 ; PRIOR FILING DATE: 2002-10-09
 ; PRIOR APPLICATION NUMBER: US 09/266.965
 ; PRIOR FILING DATE: 1999-03-12
 ; PRIOR APPLICATION NUMBER: US 08/624.447
 ; PRIOR FILING DATE: 1996-08-19
 ; PRIOR APPLICATION NUMBER: PCT/US94/11279
 ; PRIOR FILING DATE: 1994-10-06
 ; PRIOR APPLICATION NUMBER: US 08/133.963
 ; PRIOR FILING DATE: 1993-10-07
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 66
 ; LENGTH: 1056
 ; TYPE: DNA
 ; ORGANISM: Streptomyces lavendulae
 ; US-10-267-255-66

Query Match 15.1%; Score 64.4; DB 15; Length 1056;
 Best Local Similarity 51.0%; Pred. No. 5.5e-07;
 Matches 152; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 92 CCAGGCGCATTCAGAGAGACCGGCGCACCGTGGCGCGGTGGAGAGACGGCAAGACCGGC 151
 Db 683 CCGGCGCGCGGCTCGCGGACCGGTGGAATCTGCGCGGCGCATCTTTCGAGAACCATCC 742
 QY 152 CCGAGCGAGCGGACCTGCTTGGCCCGGCTGCGCCAGGAGTGGCGGCTCGACCTCGAGAG 211
 Db 743 CCGAGCGGCGCGGACGCTTACCTATTAAGACAGCTGTGACGAGCTGGAGCGAGAGAG 802
 QY 212 CCTGCGCGCGGAGGTCTGCGCGCGGCGGATCAACCGCGAGCGACCCCAACCATGAGC 271
 Db 803 TCGTAGCGATCTCTCGCGCGGATGCGACCGCATGAAGCGGAGCTCCGGCTCTGGTGA 862
 QY 272 TGGAGCGAGAAATCGAGCTGTGTCGCGACCGACCCCAAGTGAAGAGAGCATGAAGCGGC 331
 Db 863 TCGACAACTCTATCGAGCGAGCGCGCGCATGAGCGCTCTTTCGAGACCTGCTGGC 922
 QY 332 GCATATCGCCCTTAATCTGAGCGCGCGGAGCGGAGCAAGAGGCGCGGATGAGGA 389
 Db 923 TCGTCTCGTGGCGCGCGCGGAGCGCTCGAGAGCGAATTCGCGCGCTGCTGAGAA 980

RESULT 4
 US-09-953-348-76
 ; Sequence 76, Application US/09953348
 ; Publication No. US20030134398A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, David. H
 ; APPLICANT: Mao, Yingqing
 ; APPLICANT: Varoglu, Mustafa
 ; APPLICANT: He, Min
 ; APPLICANT: Sheldon, Paul
 ; TITLE OF INVENTION: MITOMYCIN BIOSYNTHETIC GENE CLUSTER
 ; FILE REFERENCE: 600.530US1
 ; CURRENT APPLICATION NUMBER: US/09/953.348

Match	Best Local Similarity	Pred. No.	Mismatches	Indels	Gaps
Matches 153;	Conservative	0;	146;	0;	0;

! PUBLICATION NO. US2004021421A1
! GENERAL INFORMATION:

```

1  APPLICANT: La Rosa, Thomas J.
2  APPLICANT: Kovalic, David K.
3  APPLICANT: Zhou, Yihua
4  APPLICANT: Cao, Yongwei
5  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
6  TITLE OF INVENTION: Plants
7  FILE REFERENCE: 38-21(53222)B
8  CURRENT APPLICATION NUMBER: US/10/425,115
9  CURRENT FILING DATE: 2003-04-28
10 NUMBER OF SEQ ID NOS: 369326
11 SEQ ID NO 73639
12 LENGTH: 683
13 TYPE: DNA
14 ORGANISM: Zea mays
15 FEATURE:
16 NAME/KEY: unsure
17 LOCATION: (1)..(683)
18 OTHER INFORMATION: unsure at all n locations
19 FEATURE:
20 OTHER INFORMATION: Clone ID: MRF4577_167158C.1
21 US-10-425-115-73639

```

Query Match	14.2%	Score	60.6	DB	18	Length	683
Best Local Similarity	51.3%	Pred.	No. 5.2e-06				
Matches	141	Conservative	0	Mismatches	134	Indels	0
						Gaps	0

QY	47	TCCTCAACGCGCCGGAGACCGGACCGAGACTGACCAAGAGCGAGTTGGCCAGCGGCATCCGAA	10
Db	320	TCCTTGGCGGGCGATTGAACGTGCTCGTGCACCCGTTACGGAGCGCGGCGCCGACGCTTGAGC	26
QY	107	AGGACCGGGCCACCGGTGCGCGGTTGGAGAGACGGCAAGAACCGGCGCGACGACGCGGACC	16
Db	260	GCGGCTTCGGCTCTGCAATGTGCGCGGGCGTTGACACGACGAGATTTTCCGACGCGCGGGGACC	20
QY	167	TCGTTGCCCGCGTGGCGCCAGAGTGTCTGGGCGCTGCACTGCACGAAAGCCCTCGCGCGCGGAG	22
Db	200	TCGTCTGGACAGTAGACACAGACAGTAGACCGTCTCTTCCGCGGCGCGCGCGGGGGGCGTCG	14
QY	227	GTCGTGGCGCCCGGCGGTCAACCCGCGACGACGACCCCAACCATGACCTGACGAGGAATTCG	28
Db	140	GGGCGCGGACAGGAAGCAGGGCTCGCGCGCGTACCGGCGGTGCGCGGTGAATGAACTTGTG	81
QY	287	AGCTGTCCGACACGAGACCCCAAGACTGAGCAGGAC	321
Db	80	ACGGTCCGGGCTCGAGGTCACTTTGGCGGAACC	46

RESULT 8
US-10-42

```

US-10-425-114-25190/c
? Sequence 25190, Application US/10425114
? Publication No. US20040034688A1
? GENERAL INFORMATION:
? APPLICANT: Liu, Jindong
? APPLICANT: Zhou, Yihua
? APPLICANT: Kovalic, David K.
? APPLICANT: Screen, Steven E.
? APPLICANT: Tabaska, Jack E.
? APPLICANT: Cao, Yongwei
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
? FILE REFERENCE: 38-21153131B
? CURRENT APPLICATION NUMBER: US/10/425,114
? CURRENT FILING DATE: 2003-04-28
? NUMBER OF SEQ ID NOS: 73128
? SEQ ID NO 25190
? LENGTH: 2129
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? OTHER INFORMATION: Clone ID: LIB3732-006-E3_FLI
US-10-425-114-25190

```

Query Match 14.2%; Score 60.6; DB 17; Length 2129;

```

- Best Local Similarity 51.3%; Pred. No. 4.5e-06;
Matches 141; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

Accession	Sequence	Position
QY	TCCTCACCAGCCCGGAGGCGCCCGGACTGACCAAGAGGAGATTGGCCAGCGGCATCCAGA	106
Db	1844 TCTTTGCCGCGGATGAAAGTGTGCGGTGACACCCGTACGGGACGGCGCCCGGACGTTGACG	1785
QY	107 AGGACCCGGGACCCGCTCGGCGCGGTGGAGAGACGGGGAAGAACCGGCGCGACGCGGACCC	166
Db	1784 GCGGCGCTTCGCGCCCGCATGTCGCGGGCGGTGACGACGACATTTCCAGACCGCCCGCGGCGC	1725
QY	167 TCGTTGCCGCGGTGAGCCAGGCTGTGCGACCTCGACCTCGACGAAGAGCCCTCGCGCGCGAG	226
Db	1724 TCGTCTGTGTGCAAGTAGACACAGCACTAAGCGTCTCTTCGCGGGGCGCCCGCGGGGGCGTTCG	1665
QY	227 GTCTGCGCCCCGCGCGTCAACCCCGCAGCGACCCCAACCATGAGACTTGACGAGGAAATCG	286
Db	1664 GGGCGCGGCAAGAAAGAGGCGTCCGCGCCCGGTACCGGCGGTGCGCCGTAGATGAATCTTGCG	1605
QY	287 AGCTGTCCGACCAAGCCCAAGGTGGACGAGAGC	321
Db	1604 ACGGTGCCGCGCTCGAGGTCCACTTTGGCGGAAGCC	1570

RESULT 9
US-10-156-761-3520

```

Sequence 3520, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, NASHAIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 3520
LENGTH: 1284
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1284)
US-10-156-761-3520

```

Query Match	13.8%	Score 58.6	DB 15	Length 1284
Best Local Similarity	49.2%	Pred. No. 1.5e-05		
Matches 154	0	Mismatches 159	Indels 0	Gaps 0
Conservative				

QY	3	GGCGCGATACGCTCTTCACCGCCCGGAGCCGGCCGGACCGCAAGCGCAAGAGCGATGGGCGAG	95
Db	198	GGCCGCTCTCCGCCGCTCATGACAGATGGCGCCCGAGCTCCTCGGCTTACGACCGCCACCGA	257
QY	96	GGCATATCCAGAAAGACCGGGGCCACCGTGGCGCGGTGGAGAGACCGCAAGAACCGGCCGA	155
Db	258	GCAGGCGCGTGAATCGACCAAGGGGATTTTGACCTTGGACCGCACCGCAACAAAGGGCTGGCT	317
QY	156	CGACGGGACCTTCGTTGGCCCGCGTGCAGCCAGGTGCTGGGCTCGACTCGACGAAAGCCT	215
Db	318	CGCGCGCAACGCGCATCTCGGGGTCTCCTCGCGCGTGCAGCACGCGCTCCGAGGCGCAG	377
QY	216	CGCGCGCGGAGTCTGGCGCCCGCGGTACCGCCGCGCAGCGCAACCCCAACCATGACCTGGGA	275
Db	378	CGACCTCCCGCTCTTTCGCTACTCTGGGGGGCGCCAGCGCGACCTCTGCTGCCCTTCCGAT	437

QY 276 CGAGGAATCGAGCTGCTCCGACCCGACCCCAAGCTGACGAGGACATGAGCGGCAT 335
 DB 438 GATGAACATCTCGAAGCGGGGCTCGACCGCCGACTCCAAAGTGACATCCAGGATTCAT 497
 QY 336 CATCGCCCTTAATC 348
 DB 498 GATCGCCCGCATC 510

RESULT 10

US-10-156-761-4348
 ; Sequence 4348, Application US/10156761
 ; Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 4348
 LENGTH: 1365
 TYPE: DNA
 ORGANISM: Streptomyces avermitilis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(1365)
 US-10-156-761-4348

Query Match 13.8%; Score 58.6; DB 15; Length 1365;
 Best Local Similarity 49.5%; Pred. No. 1.5e-05;
 Matches 151; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 38 CCGCATAGCTCTACCGCCGCGAGCGCGCGGACTGACCAAGACGAGTTGGCCAGGC 97
 DB 905 CCACCGTGCATGACGACCGCCGCGAGGTCGACGCTGACCAAGTGAACCTGGGCTTCG 964
 QY 98 GCATCCAGAAAGACCGGCGCAACCGTGGCCGAGTGGAGACGCAAGAACCGGCCGACG 157
 DB 965 CCGCGCAGACCGAAGCGGGGGCTGCTGCTCCGCTGTCGAGGAGCGGACGCGGAGAG 1024
 QY 158 ACGCGGACCTGCTGCGCGCGCTGCGCGAGGTCGCTGAGCTTCAACGAAAGCTTCG 217
 DB 1025 CCGAGTGTCTAGCGCGCGAGTTCGCGGAGTGAACGAGCGCGCCCGGACCGGCACTTCA 1084
 QY 218 CCGCGCAGGTCGCGCGCGCGCGCTGACCCCGGCGAGCGAAGCCCAACGATGACCTGGAAG 277
 DB 1085 CACCGGCGGAACTGACCGCGCGGCACTTCACTTGAACAACTGACGCGGCTGTCGCGCTCG 1144
 QY 278 AGGAATCGAGCTGCTCCGCGACCGAAGCTGGAAGGAGGAGCATGAAAGCGCGCATCA 337
 DB 1145 ACGGTTCCACCGCATCATCAACACCCGAGCGGCGCATGCTGGGCGCTCGGCCCATCA 1204
 QY 338 TCGCC 342.
 DB 1205 TCCCC 1209

RESULT 11
 US-10-156-761-1
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1

GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 1
 LENGTH: 9025608
 TYPE: DNA
 ORGANISM: Streptomyces avermitilis
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (418715)
 OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1

Query Match 13.8%; Score 58.6; DB 15; Length 9025608;
 Best Local Similarity 49.5%; Pred. No. 5.3e-06;
 Matches 151; Conservative 0; Mismatches 154; Indels 0; Gap

QY 38 CCGCATAGCTCTACCGCCGCGAGCGCGCGGACTGACCAAGACGAGTTGGCCAGGC 97
 DB 5342891 CCAACCGTGCATGACGACCGCCGCGAGGTCGACGCTGACCAAGTGAACCTGGGCTTCG 964
 QY 98 GCATCCAGAAAGACCGGCGCAACCGTGGCCGAGTGGAGACGCAAGAACCGGCCGACG 157
 DB 5342951 CCGCGCAGACCGAAGCGGGGGCTGCTGCTCCGCTGTCGAGGAGCGGACGCGGAGAG 1024
 QY 158 ACGCGGACCTGCTGCGCGCGCTGCGCGAGGTCGCTGAGCTTCAACGAAAGCTTCG 217
 DB 5343011 CCGAGTGTCTAGCGCGCGAGTTCGCGGAGTGAACGAGCGCGCCCGGACCGGCACTTCA 1084
 QY 218 CCGCGCAGGTCGCGCGCGCGCGCTGACCCCGGCGAGCGAAGCCCAACGATGACCTGGAAG 277
 DB 5343071 CACCGGCGGAACTGACCGCGCGGCACTTCACTTGAACAACTGACGCGGCTGTCGCGCTCG 1144
 QY 278 AGGAATCGAGCTGCTCCGCGACCGAAGCTGGAAGGAGGAGCATGAAAGCGCGCATCA 337
 DB 5343131 ACGGTTCCACCGCATCATCAACACCCGAGCGGCGCATGCTGGGCGCTCGGCCCATCA 1204
 QY 338 TCGCC 342.
 DB 5343191 TCCCC 5343195

RESULT 12
 US-10-156-761-1/c
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089

Mon Apr 18 09:47:30 2005

us-09-855-340b-2.rnpb

Db 732 GCGGTTGAACTTCGCCGTGAGCGGCCACCGGGGCGCTCGACGTTTCTACAGCAACGCCG 673

Qy 285 CGAGCTGTCCGACACCGAACCCAAAGCTGGAAGAGACATGACAGCGCGCATCACCCT 344

Db 672 CGTCTTGAGGCTCATATCGGCGGAGCGCGCTCGCTTCCTTGAGACTTGGGCGAGTTCCACG 613

Qy 345 AATCTCGAGACCGCCGTGACCGCGACACAGCGCGCGCGATCGAGGAAACCAAGCGGCTCAT 404

Db 612 CGTCAATGGCCGTGAAGCGCCCGCGCGCCGTTCGCCGCGCCAGCAGCGGCGCGCCCAT 553

Qy 405 CGACCTGTTCCGCGGAGCTG 425

Db 552 GGTGCGCGCGCCGGAAGCGGCTG 532

```

Oy      364  CGCAGCAAGGGGGGGCGATC 384
          |||||
Db      1090  GTCGCACTGGTGATGATC 1110
          |||||

```

Search completed: April 15, 2005, 22:51:17
Job time : 339.016 secs

```

RESULT 5
US-10-156-761-6116
; Sequence 6116, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6116
LENGTH: 1659
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1659)
US-10-156-761-6116

```

Query Match	13.6%	Score 57.8;	DB 15;	Length 1659;
Best Local Similarity	47.0%;	Pred. No. 2.3e-05;		
Matches 179; Conservative	0;	Mismatches 202;	Indels 0;	Gaps 0;

QY 4 GGCAACACCGGCGGCTGGGGCGGGGCAATGGGCGCATGCTCTCACCGCCCGCGAG 63
 Db 730 CACAAAGGCGCGCGTGTCCGATCTGGGACGACCACGAGTTCTGGCGCAACGCGGTGTGG 789
 QY 64 CGCCCGCGCATGACCAAGAGGAGTGTGGCCAGCGCATCTCAAGAAAGGACCGGGCCACCGTTC 123
 Db 790 GGCAGGCGGGGAAACACACGAGGGCGCCCGAAGGCACTTGGACGGCACGTCAGAGCGCC 849
 QY 124 GGCAGGTGGAGGACCGGCAGAAACCGGCGCCACGACGACGTCCTGTGTCCCGCGTCCGC 183
 Db 850 GCGAAACAGGCGCTACTTTCAGATGGATGTCGCGGTCCGCCCGCGGATCGAAGGGACCACTTAC 909
 QY 184 CAGGTGCTCGGCTCTCGACCTTCGACGGAACCCCTCGCGCCGCGCAGGTCTTGCGCCCGGCGTTC 243
 Db 910 CGGCGGCTGCGCTTCGGCAAGCTCGCGACCTCTGTGCTCGACCTTGGCGGTCTTCCGC 969
 QY 244 ACCCGCCGAGGACCCCAACCATGACCTTGGACGAGGAATCGAGCTGTGTCGACCGAC 303
 Db 970 TCGCAGCAGGGGTCCACGAGGACGCGCACGCGTCGACGCCGAGGCGTACCTCAAGGGGC 1029
 QY 304 CCCAAGCTGACGAGGACATGAAGCGGCGCATTCATCGCCCTTAATCTTGGAGCGCCGTGAG 363
 Db 1030 CGGCTCAGCTCGACTGCTGAAGCGGGGCTCAAGAGGTCTCGACACCACTGCGGGCTG 1089

THIS PAGE BLANK (USPTO)


```
XX WPI: 2002-082983/11.
DR P-PSDB; AAE15909.
XX
XX Novel polynucleotides encoding integrase, excisionase and an integrase
PT attachment site isolated from a lysogenic phage pMLP1, useful for
XX transforming an actinomycete.
PS Claim 23; Fig 4; 34pp; English.
XX
XX The present invention relates to novel polynucleotides encoding integrase
CC (int) and excisionase (xis) and an integrase attachment site (attP) which
CC are isolated from pMLP1, a bacteriophage (lysogenic phage) isolated from
CC Micromonospora carbonacea var. africana. Polynucleotides of the invention
CC are useful for transforming an actinomycete with a vector. They are also
CC useful for creating vectors for site-specific integration into host
CC chromosomes. The integrating vectors are used to express actinomycete
CC genes, manipulate secondary metabolic pathways and create new metabolic
CC products such as hybrid antibiotics. The present sequence is pMLP1
CC attP/attB right junction DNA from Micromonospora carbonacea
XX
SQ Sequence 247 BP; 38 A; 82 C; 90 G; 37 T; 0 U; 0 Other;
Query Match 100.0%; Score 247; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 1e-53;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAGGGGAATCCACTCCGAGACGCCCGGACATCCGAGCATGACGAGCAACGACG 60
DB 1 TAGGGGAATCCACTCCGAGACGCCCGGACATCCGAGCATGACGAGCAACGACG 60
QY 61 GTCAGGTGGCCCTGTGACCCCTGACACGAGGCCCGGTACGGGTTCAATTCCTCATGTC 120
DB 61 GTCAGGTGGCCCTGTGACCCCTGACACGAGGCCCGGTACGGGTTCAATTCCTCATGTC 120
QY 121 ACCCGTACACGAGAGGCCCGCTTCCATCCGAGAGGGGCGCTTCTGAGGGTTCCGCG 180
DB 121 ACCCGTACACGAGAGGCCCGCTTCCATCCGAGAGGGGCGCTTCTGAGGGTTCCGCG 180
QY 181 GTCAGGCGGTGGCTCGGCGCTGAGGAGACTCGGCCCGGCGGAGGTGGCTCGGCGT 240
DB 181 GTCAGGCGGTGGCTCGGCGCTGAGGAGACTCGGCCCGGCGGAGGTGGCTCGGCGT 240
QY 241 CCGGGGA 247
DB 241 CCGGGGA 247
RESULT 2
AAD25934
ID AAD25934 standard; DNA; 241 BP.
XX
XX AAD25934;
AC
XX 26-MAR-2002 (first entry)
DT
XX
XX Micromonospora carbonacea pMLP1 attB DNA.
DE
XX
XX Integrase; int; excisionase; xis; integrase attachment site; attP; pMLP1;
KW site-specific integration; hybrid antibiotic; metabolic product; attB;
XX
XX Micromonospora carbonacea.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 44..119
FT /*tag= a
FT /*note= "rRNA-His DNA"
FT misc_feature 95..119
FT /*tag= b
FT /*note= "attB region"
FT repeat_region 124..137
FT /*tag= c
```

```
FT /rpt_type= INVERTED
FT /note= "Inverted repeat 1 (IR1)"
FT repeat_region 142..155
FT /*tag= d
FT /rpt_type= INVERTED
FT /note= "Inverted repeat 2 (IR2)"
FT CDS complement(179..241)
FT /*tag= e
FT /product= "attB peptide"
XX
XX WO200187936-A2.
XX
XX 22-NOV-2001.
XX
XX 15-MAY-2001; 2001WO-US015760.
XX
XX 17-MAY-2000; 2000US-0204670P.
XX
XX (SCHE ) SCHERING CORP
XX
XX Hosted TJ, Horan AC;
XX
XX WPI: 2002-082983/11.
XX P-PSDB; AAE15909.
XX
XX Novel polynucleotides encoding integrase, excisionase and an integrase
PT attachment site isolated from a lysogenic phage pMLP1, useful for
XX transforming an actinomycete.
PS Claim 23; Fig 4; 34pp; English.
XX
XX The present invention relates to novel polynucleotides encoding integrase
CC (int) and excisionase (xis) and an integrase attachment site (attP) which
CC are isolated from pMLP1, a bacteriophage (lysogenic phage) isolated from
CC Micromonospora carbonacea var. africana. Polynucleotides of the invention
CC are useful for transforming an actinomycete with a vector. They are also
CC useful for creating vectors for site-specific integration into host
CC chromosomes. The integrating vectors are used to express actinomycete
CC genes, manipulate secondary metabolic pathways and create new metabolic
CC products such as hybrid antibiotics. The present sequence is pMLP1 attB
CC DNA from Micromonospora carbonacea
XX
SQ Sequence 241 BP; 27 A; 78 C; 92 G; 44 T; 0 U; 0 Other;
Query Match 59.5%; Score 147; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 3.7e-28;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 GGGTTCAATTCATCATGATCAACCGGTACAGAGGCCCTTCCACTGAGAGGGGCTTC 160
DB 95 GGGTTCAATTCATCATGATCAACCGGTACAGAGGCCCTTCCACTGAGAGGGGCTTC 154
QY 161 GGGCTTCAGAGGGTTCGCGGTAGAGGGCTGCGGCTGAGGAGCTCGGAGGAGCTCGGCGT 220
DB 155 GGGCTTCAGAGGGTTCGCGGTAGAGGGCTGCGGCTGAGGAGCTCGGAGGAGCTCGGCGT 214
QY 221 GCGGAGAGTGGCTCGGCGCTCGGAGG 247
DB 215 GCGGAGAGTGGCTCGGCGCTCGGAGG 241
RESULT 3
AAD25945
ID AAD25945 standard; DNA; 209 BP.
XX
XX AAD25945;
AC
XX 26-MAR-2002 (first entry)
DT
XX
XX Micromonospora carbonacea pMLP1 attP DNA.
DE
XX
XX Integrase; int; excisionase; xis; integrase attachment site; attP; pMLP1;
KW site-specific integration; hybrid antibiotic; metabolic product;
```

DT	11-SEP-2003	(revised)	
DT	26-SEP-2001	(first entry)	
XX			
DE	Micromonospora DNA encoding integrase enzymes.		
KM	Evernimycin; antibiotic; bottle-neck gene; orthomycin; fermentative		
XX	integrase; ds.		
OS	Micromonospora sp. ATCC 39149.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	963..1387	
FT		/*tag= a	
FT		/product= "Integrase #1"	
FT	CDS	1394..2572	
FT		/*tag= b	
FT		/product= "Integrase #2"	
FT		2570..2799	
FT	misc_feature	/*tag= c	
FT		/note= "AttC/B/AttP region of integrase action"	
FT		2714..2715	
FT	misc_recomb	/*tag= d	
FT		/label= Insertion junction	
FT		/note= "Site of integrase activity"	
XX			
PN	WO200151639-A2.		
XX			
XX	19-JUL-2001.		
XX			
PD	12-JAN-2001; 2001WO-US001187.		
XX			
PF	12-JAN-2000; 2000US-0175751P.		
PR			
XX			
PA	(SCHE) SCHERING CORP.		
PI	Hosted TJ, Horan AC, Wang TX;		
XX			
DR	WPI, 2001-442147/47		
XX	P-PSDB; AAU04900, AAU04912.		
PT	New nucleic acid molecules encoding evernimycin pathway gene product		
PT	useful for improving yields of evernimycin, to produce new		
XX	evernimycin and as probes to identify homologous sequences.		
XX			
PS	Claim 26; Fig 7; 109pp; English.		
XX			
CC	The sequence encodes 2 integrase which permit site specific integ		
CC	Of a vector into an actinomycete, especially a Micromonospora, gen		
CC	The invention relates to nucleic acids and vectors comprising a M.		
CC	carbonacea evernimycin biosynthetic pathway resistance gene prod		
CC	useful for selecting for a transfected or transformed host cell. A		
CC	integrative version of the vector is useful for introducing a		
CC	evernimycin pathway gene (a bottle-neck gene) into an actinomyc		
CC	the genus Micromonospora. The DNA encoding the biosynthetic protei		
CC	useful for synthesising novel evernimycin-related compounds, ari		
CC	from modifications of the DNA sequence designed to change glycosyl		
CC	modified orsellinic acid groups contained in evernimycin, for		
CC	expressing functional or mutant evernimycin biosynthetic enzyme		
CC	evaluation, diagnosis and preferably biosynthesis of evernimycin		
CC	other secondary metabolic products, improving the yield of evernim		
CC	and to produce novel evernimycins and also as a hybridisation pro		
CC	identify homologous sequences. The encoded polypeptides are useful		
CC	combinatorial biosynthesis to generate libraries of orthonycins, e		
CC	evernimycin analogues/homologues and drug discovery. The DNA enc		
CC	the integrase allows for increasing a given gene dosage. The integ		
CC	vector can be used to permanently integrate copies of a heterologous		
CC	of choice into chromosomes of different hosts and to integrate gen		
CC	which increase the yield of known products or to generate novel pr		
CC	such as hybrid antibiotics or other novel secondary metabolites. T		
CC	vector can also be used to integrate antibiotic resistance genes i		
CC	to carry out bioconversions with compounds to which the strain is		
CC	normally sensitive and is thus useful in fermentation processes in		
CC	e.g. Streptomyces antibiotics. (Updated on 11-SEP-2003 to standar		

PS	Claim 23; Fig 5; 34pp; English.
XX	
XX	The present invention relates to novel polynucleotides encoding integrase
CC	(int) and excisionase (xis) and an integrase attachment site (attP) which
CC	are isolated from pMPL1, a bacteriophage (lysogenic phage) isolated from
CC	Micromonospora carbonacea var. africana. Polynucleotides of the invention
CC	are useful for transforming an actinomycete with a vector. They are also
CC	useful for creating vectors for site-specific integration into host
CC	chromosomes. The integrating vectors are used to express actinomycete
CC	genes, manipulate secondary metabolic pathways and create new metabolic
CC	products such as hybrid antibiotics. The present sequence is pMPL1
CC	attP/attB right junctionure DNA from Micromonospora halophilica. (Updated on
CC	07-APR-2003 to correct OS field.)
XX	
SQ	Sequence 260 BP; 47 A; 83 C; 90 G; 40 T; 0 U; 0 Other;
Query Match	50.2%; Score 124; DB 6; Length 260;
Best Local Similarity	100.0%; Pred. No. 2,8e-22;
Matches 124; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
Qy	1 TGGGGGAATCCCATCTCCGGAGAGCCCGGAGCAATCCGGAGCATGACGAGCAACCCAGCAG 60
Db	1 TGGGGGAATCCCATCTCCGGAGAGCCCGGAGCAATCCGGAGCATGACGAGCAACCCAGCAG 60
Qy	61 GTCAGATGAGCCCTGTGACCCCTGACCAAGGGCCCGGTACGGGTTCAATTCCTCATGTC 120
Db	61 GTCAGATGAGCCCTGTGACCCCTGACCAAGGGCCCGGTACGGGTTCAATTCCTCATGTC 120
Qy	121 ACCC 124
Db	121 ACCC 124
RESULT 6	
ADT44391/c	
ID	ADT44391 standard; cDNA; 1446 BP.
XX	
AC	ADT44391;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Bacterial polynucleotide #19142.
XX	
KW	Recombinant DNA construct; transformed plant; improved plant property;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW	pathogen tolerance; pest tolerance; plant disease resistance;
KW	cell cycle pathway modification; plant growth regulator;
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW	bacterial polynucleotide; gene; ss.
XX	
OS	Bacteria.
XX	
FN	US2003233675-A1.
XX	
PD	18-DEC-2003.
XX	
PF	20-FEB-2003; 2003US-00369493.
XX	
PR	21-FEB-2002; 2002US-0360039P.
XX	
PA	(CAOV/) CAO Y.
PA	(HINK/) HINKLE G J.
PA	(SLAT/) SLATER S C.
PA	(CHEN/) CHEN X.
PA	(GOLD/) GOLDMAN B S.
XX	
PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	
DR	WPI; 2004-061375/06.
XX	

1 The invention relates to a recombinant DNA construct comprising a
2 promoter functional in a plant cell, where the promoter is positioned to
3 provide for expression of a polynucleotide encoding a polypeptide from a
4 microbial source. The invention also relates to a transformed plant
5 comprising the recombinant DNA construct and a method of producing a
6 transformed plant having an improved property. The plant is a crop plant
7 such as maize or soybean. The method of producing a plant with the
8 having an improved property comprises transforming a plant with the
9 recombinant DNA construct and growing the transformed plant, where the
10 polynucleotide or polypeptide is useful for improving plant properties.
11 The recombinant DNA construct is useful for producing plants with
12 improved plant properties, e.g. improved cold, heat or drought tolerance,
13 tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
14 increased resistance to plant disease, better growth rate by modification
15 of the cell cycle pathway with plant growth regulators, increased rate of
16 homologous recombination, modified seed oil or protein yield and/or
17 content, improved yield by modification of carbohydrate, nitrogen or
18 phosphorus use and/or uptake, by modification of photosynthesis or by
19 providing improved plant growth and development under at least one stress
20 condition, improved lignin production or improved galactomannan
21 production. This sequence represents a bacterial polynucleotide used in
22 the scope of the invention. Note: The sequence data for this patent did
23 not form part of the printed specification but was obtained in electronic
24 format from USPTO at seqdata.uspto.gov/sequence.html.

```

QY      133 AGGCCCCCTCACTCGAGGGGCGCTTCGCGGTTCGAGGGTCCGCGGTCAAGCGGTCG   192
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1227 AAGCGTTCTCGGCGCAGCTGGGCTTCGCCGTCCGAGGCCACACCGGTCCGGCGCGGG   1168

```

Dy Db

	193	GCTCGGCGCTGGGGGA	ACTCGGCCCGTCCGCGGGAGTGAGCCTCGGCGGTCCG	244
	1167	GAGCGCGCGCGGCAC	CTTCCGTTCCGTCGAGCGGCAAGGCTCCAG	1116

RESULT 7.
AAZ29447/c
ID AAZ29447 standard; DNA; 3396 BP.

DT 14-MAR-2000 (first entry)

DNA-2 related to human gb2 GABA B receptor.

gβ2 GABA B receptor; human gβ2; Hgb2; calcium-sensing receptor;
KW gamma- α -methyl glutamic acid; 4-amino butanoic acid; GABA;
KW metabotropic glutamate receptor; neurological disorder;
KW psychiatric disorder; agonist, antagonist, ss.

OS · Homo sapiens.

Key	Location/Qualifiers
CDs	458. .3355

```

Ft      /csg-
Ft      /transl except=(pos: 755..760, aa: Asp)
Ft      /note="The amino acid residue corresponding to
Ft      758 to 760 is not given in the specification"

```

PN	WO9961606-A1
XX	
PD	02-DEC-1999.
XX	
PF	28-MAY-1999;

28-MAY-1999; 99WO-US011869.

XX
PR 29-MAY-1998; 98US-0087274P.

AA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA

Clark J, Bonner TI;

DR WPI; 2000-105616/09.

DR P-PSDB; AAY44345.

AA New GABA_B (c-aminobutyric acid or 4-aminobutanoic acid) receptor, used
PT for identification of (ant)agonists and for treatment of neurological
PT disorders.

PS Disclosure; Page 58; 67pp; English.

The present sequence is a DNA related to human gb2 GABA B receptor. gb2 (Hgb2) shares sequence homology with rat GABA B receptor, rgb1, parathyroid cell calcium-sensing receptor and metabotropic glutamate receptors. It can be produced in host cells by transforming them with recombinant expression vector comprising gb2 encoding cDNA. The cell that express the receptor are used in the development of drugs for treatment of neurological and psychiatric disorders, for pharmacological, functional, or other investigational analysis of gb2 physiological, functional, or other investigational analysis of the GABA B receptor, its agonists or antagonists and for determining the ability of a chemical to bind to a mammalian gb2 GABA B receptor *in vitro*. It may also be used for the preparation of antibodies to gb2 which can be used in diagnostic assays. Note: There is no relevant information about this sequence in the specification

Sequence 3396 BP; 761 A; 1062 C; 940 G; 633 T; 0 U; 0 Other;

Query Match	15.5%	Score 38.2;	DB 3;	Length 3396;
Best Local Similarity	47.3%;	Pred. No. 3.1;		
Matches 115;	Conservative 0;	Mismatches 128;	Indels 0;	Gap

Dy
Dd

3 GGGGAAATTCACCTCCGGAGACGCCCGGAGCAATCCGAGAGCATGAACGAGAACCAACCGAGAAG
| | | | | | | | | | | | | | | | | | | | | |
512 CGGGTGGCGGCGCGAGCGCGCGCGCCCGGGCTGCCTCGAGACTCCGCGGGGGAAGCATTGCC

Dy 63 CAGGTGACCTGTTGACCCCTTACACAGGGCCCCGGTACGGGTTCAATTCCCATCAGTC
| | | | |
Db 452 GCCCGGAGTGCCGGGCCTCGGCTACTTCGGCCCGCATGGCCTTGGCCCGGCCCGCGGCC

Db

392 GCGCCAGCTCTTCGCGCGCGCGCGCAATGCGCGCGCGCGCGCGCGCGCTCGCG

100

— — — — —

Db 272 GTG 270

RESULT 8
AAZ29446/C

XXI

DT 14-MAR-2000 (first entry)

AA
DE DNA-1 related to human gb2 GABA B receptor.

XX gb2 GABA B receptor; human gb2; Hgb2; calcium-sensing receptor;
KW gamma-amino butyric acid; 4-amino butanoic acid; GABA;
KW metabotropic glutamate receptor; neurological disorder;
KW psychiatric disorder; agonist; antagonist; ss.

OS Homo sapiens.

XX	AA229422;
AC	
XX	14-MAR-2000 (first entry)
DT	
XX	
DE	Human gb2 GABA B receptor encoding cDNA.
XX	
KW	gb2 GABA B receptor; human gb2; Hgb2; calcium-sensing receptor;
KM	gamma-amino butyric acid; 4-amino butanoic acid; GABA,
KV	metabotropic glutamate receptor; neurological disorder;
XX	psychiatric disorder; agonist; antagonist; ss.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS /tag= a
FT	/product= "Human gb2 GABA B receptor"
PX	MO961606-AI.
PN	
PD	02-DEC-1999.
XX	
PF	28-MAY-1999; 99WO-US011869.
PR	29-MAY-1998; 98US-0087274P.
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	
PI	Clark J, Bonner TI;
DR	WPI, 2000-105616/09.
XX	P-PSDB; AAY44342.
PT	New GABAB (c-aminobutyric acid or 4-aminobutanoic acid) receptor, useful
PT	for identification of (ant)agonists and for treatment of neurological
PT	disorders.
XX	
XX	Claim 4; Page 21-23; 67pp; English.
PS	
XX	
CC	The present sequence is a cDNA obtained from human brain cDNA library. It
CC	encodes gb2 subunit of GABA B receptor. Human gb2 (Hgb2) shares sequence
CC	homology with rat GABA B receptor rgb1, parathyroid cell calcium-sensing
CC	receptor and metabotropic glutamate receptors. It can be produced in host
CC	cells by transforming them with recombinant expression vector comprising
CC	hgb2 encoding cDNA. The cells that express the receptor are used in the
CC	development of drugs for treatment of neurological, functional, or other
CC	disorders, for pharmacological, physiological, and psychiatric
CC	investigational analysis of gb2 GABA B receptor, its agonists or
CC	antagonists and for determining the ability of a chemical to bind to a
CC	mammalian gb2 GABA B receptor in vitro. They may also be used for the
CC	preparation of antibodies to hgb2 which can be used in diagnostic assays
XX	
XX	
SQ	Sequence 5786 BP, 1415 A; 1670 C; 1474 G; 1227 T; 0 U; 0 Other;
	Query Match 15.5%; Score 38.2; DB 3; Length 5786;
	Best Local Similarity 47.3%; Pred. No. 3.3;
	Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0
OY	3 GGGGAATTCACCTCCGAGAGCGCCGGAGCAATCCGAGCATGACGAGCAACAGCAGT 62
Db	512 GC GGTTGC CGCGCGCGCGCGGCCCGGCGCTCCGCGAGCTCCGCGGAGACCATGCCGC 453
OY	63 CAGGTGCGCTGTGACCCTTGACACAGGCCCCCGGTACGGGTTCAATTCCATGATCAC 122
Db	452 GCCCGCGGCTGCAGCGCGCGCTGACTCGGCTCCGATGAGCCTGCGCCGCGCCGCCCCC 393
OY	123 CCGTAACAAGAAGCCCCCTTCATCTCGAAGGGGGCTTTCGGGCTTCTTAGAGGTTGCGGCT 182
Db	392 GCGCCAAGCTTTTCCGCGCGCGCGCGCAATGAGCGCCGCGGCCCGCGCTCCGCT 333
OY	183 CAGGCGGTGCGCTCGCGCGCTGCGGAGACTCGGCCCTGCGCGGAGATGCGCTCGGCGTCC 242

The present invention describes an isolated nucleic acid molecule, which comprises the sequence of any of the genes that are up-regulated or down-regulated in specific cancers (e.g. about 1031 genes up-regulated in acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer-related gene nucleotide sequences that encode the proteins given in ABR8551 to ABR8709. Also described: (1) determining the presence or absence of a pathological cell in a patient; (2) an expression vector comprising a nucleic acid molecule described above; (3) a host cell comprising the vector; (4) an isolated polypeptide, which is encoded by the nucleic acid; (5) an antibody that specifically binds the polypeptide of (4); (6) specifically targeting a compound to a pathological cell in a patient by administering to the patient the antibody above; and (7) a drug screening assay. The nucleic acid is useful as diagnostic markers or therapeutic targets. In particular the nucleic acid is useful for diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow, bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary, pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,

XX

of high nutritional value.

XX Claim 25; SEQ ID NO 1271; 230bp; English.

PS
XX
CC The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 1239 BP; 283 A; 310 C; 372 G; 274 T; 0 U; 0 Other;

Query Match 15.0%; Score 37; DB 12; Length 1239;

Best Local Similarity 57.3%; Pred. No. 5.7;

Matches 67; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY	128	CACGAAAGGCCCCCTCCACATCGAGAGGGGCGCTTCGGCGCTTCCTGAGGGTTCCGGATCAGGC	187
DB	228	CCCGAAGCGCGGCTTCACACGAGAGGCCACCGGTGTCCCGCGGGGTCAAGCCGTGGCGAC	169
QY	188	GGTCGGCTCGCGCTCGGGGACTCGGCCCGCTCGCGCGAGTGGGCTTCGGCGTCCGG	244
DB	168	GGTGAACCGCGCGCGCGCGGCGGCTTGAGCAGTCCGCGCGGCGCTTCACCCG	112

Search completed: April 15, 2005, 16:13:40
Job time : 461 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	247	100.0	247	6	AX338975	AX338975 Sequence
2	147	59.5	241	6	AX338973	AX338973 Sequence
3	145	58.7	145	1	AY150033	AY150033 Micromono
4	125	50.6	2025	7	AY150027	AY150027 Bacterio
5	124	50.2	260	6	AX338978	AX338978 Sequence
6	90	36.4	145	1	AY150028	AY150028 Micromono
7	79	32.0	143	1	AY150032	AY150032 Micromono
8	43	17.4	33064	2	AY338477	AY338477 Streptomy
9	42.6	17.2	175217	10	AC079680	AC079680 Mus muscu
10	42.6	17.2	194452	2	AC115464	AC115464 Mus muscu
11	42.4	17.2	1976	1	SEABCT	SEABCT
12	41.4	16.8	36654	9	AF315943	AF315943 Homo sapi
13	41.4	16.8	55590	9	AC073128	AC073128 Homo sapi
14	41.4	16.8	152129	2	AC027416	AC027416 Homo sapi
15	41.4	16.8	310550	1	SC0939113	AL939113 Streptomyce
16	41.2	16.7	5966	1	SMR118862	Y18862 Streptomyce
17	41	16.6	301443	1	AE017239	AE017239 Mycobacte
18	40.8	16.5	600	1	MUTRNAGC	AE017239 Mycobacte
19	40.8	16.5	163386	9	AC015723	AE015723 Homo sapi

C	20	40.2	16.3	300800	1	SC0939112
C	21	40	16.2	300425	1	AP005056
C	22	39.8	16.1	5973	1	AB024038
C	23	39.8	16.1	6573	1	AB027763
C	24	39.8	16.1	27680	1	SC0939115
C	25	39.8	16.1	300425	1	AP005056
C	26	39.4	16.0	349497	1	BK640440
C	27	39.2	15.9	124581	9	BK643634
C	28	39.2	15.9	170935	2	AC016079
C	29	39.2	15.9	302300	1	AP0050534
C	30	39	15.8	110000	1	AP006840
C	31	39	15.8	349116	1	AP003003
C	32	38.8	15.7	300327	1	AE017228
C	33	38.6	15.6	38911	9	AC005559
C	34	38.4	15.5	2070	5	AP2457675
C	35	38.4	15.5	110000	1	BK571866
C	36	38.2	15.5	177166	2	AC123379
C	37	38.2	15.5	5786	9	AF056085
C	38	38.2	15.5	13219	1	AE005106
C	39	38	15.5	153099	2	AC1498983
C	40	38	15.4	308050	1	SC0939112
C	41	37.8	15.3	302675	1	AP0050524
C	42	37.6	15.2	84811	9	HS469922
C	43	37.6	15.2	125020	9	AF429315
C	44	37.6	15.2	186829	9	BK400519
C	45	37.4	15.1	791	5	CR733014

AL9391128 StrepD
 AP0050038 StrepD
 AB0240536 StrepE
 AB0217763 StrepE
 AL9391151 StrepE
 AP0050044 StrepE
 BX6050040 Bordet
 BX6493364 Humana
 AC0160790 Homo
 AP0050034 StrepE
 Continuation
 Continuation
 AP0030023 Mesoro
 AE011228 Mycol
 AC0055556 Homo
 AF2461676 Gallu
 Continuation
 AC123979 Papi
 AF056085 Homo
 AE005106 Halo
 AC1499873 Strep
 AL9391124 Strep
 AP0050524 Strep
 AL0311284 Humana
 AF429515 Homo
 BX6405019 Humana
 CR733014 Gallu

RESULT 1	247 bp	DNA	linear	PAT 09-0
AX338975				
LOCUS				
DEFINITION	Sequence 6 from Patent WO0187936.			
ACCESSION	AX338975			
VERSION	AX338975.1	GI:18129111		
KEYWORDS				
SOURCE				
ORGANISM	Micromonospora carbonacea			
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales,			
	Micromonosporinae; Micromonosporaceae; Micromonospora.			
REFERENCE				
AUTHORS	1 Hosted, T. J. and Horan, A. C.			
TITLE	Isolation of Micromonospora carbonacea var africana pmpl1 and use of integrating function for site-specific integration and use of Micromonospora carbonacea ch			
JOURNAL	Patent: WO 0187936-A 6 22-NOV-2001;			
	SCHERING CORPORATION (US)			
FEATURES				
source	1 . 247			
	/organism="Micromonospora carbonacea"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:47853"			
ORIGIN				
	100.0%; Score 247; DB 6; Length 247;			
Query Match	Best Local Similarity 100.0%;			
Matches	247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 TAGGGGATTCACATCCGGAGACGCCCGAGACCAATCCGGAGCATGACGAGCAACGAG			
Db	1 TAGGGGAAATCACTCTCCGAGACGCCCGAGACCAATCCGGAGCATGACGAGCAACGAG			
QY	61 GTGAGATGGCTGTGTGACCCCTTGACGAGGCCCCGGTACGGTTCAATTCCATCA			
Db	61 GTGAGATGGCTGTGTGACCCCTTGACGAGGCCCCGGTACGGTTCAATTCCATCA			
QY	121 ACCGTACACGAAAGCCCCCTTCACCTGGAGAGGGGGACCTTGCGGGTTCCTGAGGGTTC			
Db	121 ACCGTACACGAAAGCCCCCTTCACCTGGAGAGGGGGACCTTGCGGGTTCCTGAGGGTTC			
QY	181 GTGAGGCGGTGGGCTCGGCGCTGGGAGGACTTGAGGCCCGTGGCGGGAGTGCGCTCGG			

Db 181 GTGAGGCGGTGGGCTGGCCCTGGGGACTCGGCCCTGGGAGTGGCTTCGGCGT 240
Qy 241 CCGGGGA 247
Db 241 CCGGGGA 247

RESULT 2
AX338973
LOCUS AX338973 241 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 4 from Patent WO0187926.
ACCESSION AX338973
VERSION AX338973.1 GI:18129109
KEYWORDS
SOURCE Micromonospora carbonacea
ORGANISM Micromonospora carbonacea
Bacteria; Actinobacteria; Actinomycetales;
Micromonosporinae; Micromonosporaceae; Micromonospora.

REFERENCE
AUTHORS Hosted,T.J. and Horan,A.C.
TITLE Isolation of Micromonospora carbonacea var africana pmlp1 integrase
JOURNAL and use of integrating function for site-specific integration into
Patent: WO 0187936-A 4 22-NOV-2001;
SHERING CORPORATION (US)
FEATURES
source Location/Qualifiers
1..241
/organism="Micromonospora carbonacea"
/mol_type="unassigned DNA"
/db_xref="taxon:47853"

ORIGIN
Query Match 59.5%; Score 147; DB 6; Length 241;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 GGGTTCATTCCTCCATCAGTACCGGTACAGAGGCCCCCTTCCTCGAGGGGCGCTTC 160
Db 95 GGGTTCATTCCTCCATCAGTACCGGTACAGAGGCCCCCTTCCTCGAGGGGCGCTTC 154
Qy 161 GGGCTTCCTGAGGGGCTCGCGGTACAGGCGGTGGCTGGGGGACTCGGCGCCGCTC 220
Db 155 GGGCTTCCTGAGGGGCTCGCGGTACAGGCGGTGGCTGGGGGACTCGGCGCCGCTC 214
Qy 221 GGGCGAGTGGCTCGCGGTGGGGGA 247
Db 215 GGGCGAGTGGCTCGCGGTGGGGGA 241

RESULT 3
AY150033 145 bp DNA linear BCT 02-SEP-2003
LOCUS AY150033
DEFINITION Micromonospora sp. ATCC 39149 attR region.
ACCESSION AY150033
VERSION AY150033.1 GI:28630429
KEYWORDS
SOURCE Micromonospora sp. ATCC 39149
ORGANISM Micromonospora sp. ATCC 39149
Bacteria; Actinobacteria; Actinomycetales;
Micromonosporinae; Micromonosporaceae; Micromonospora.

REFERENCE
AUTHORS Alexander,D.C., Devlin,D.J., Hewitt,D.D., Horan,A.C. and
Hosted,T.J.
TITLE Development of the Micromonospora carbonacea var. africana ATCC
JOURNAL 39149 bacteriophage pMLP1 integrase for site-specific integration
PUBMED in Micromonospora spp
12949170 Microbiology (Reading, Engl.) 149 (Pt 9), 2443-2453 (2003)
REFERENCE 2 (bases 1 to 145)
AUTHORS Hosted,T.J. Jr., Hewitt,D.D. and Alexander,D.C.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2002) New Lead Discovery, Schering-Plough

FEATURES Research, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA
source Location/Qualifiers
56..145
/organism="Micromonospora sp. ATCC 39149"
/mol_type="genomic DNA"
/strain="ATCC 39149"
/db_xref="ATCC:39149"
/db_xref="taxon:219305"
/focus
/note="deposited in ATCC as Micromonospora carbonacea var.
africana Waltz et al."
1..55
/organism="Integration vector pSPRH840"
/mol_type="genomic DNA"
/db_xref="taxon:219291"
1..145
/note="attR region"
55..56
/note="site of pSPRH840 recombination"
56..79
/note="pSPRH840 attachment site; attB/attP identity
element"
join(85..98,103..116)

ORIGIN
stem_loop
Query Match 58.7%; Score 145; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 7.7e-18;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 CGAGCAACGACGAGTGGCTGTGACCCCTGACACGAGGCCCCGCTACGGGTT 105
Db 1 CGAGCAACGACGAGTGGCTGTGACCCCTGACACGAGGCCCCGCTACGGGTT 60
Qy 106 CAATTCCTATGACACCCGTAACGAGGCCCCCTTCCTCGAGGGGCTTCGGCGT 165
Db 61 CAATTCCTATGACACCCGTAACGAGGCCCCCTTCCTCGAGGGGCTTCGGCGT 120
Qy 166 TCGTGGGTTTCGGGTGAGGCGGT 190
Db 121 TCGTGGGTTTCGGGTGAGGCGGT 145

RESULT 4
AY150027 2025 bp DNA linear PHG 02-SEP-2003
LOCUS AY150027
DEFINITION Bacteriophage pMLP1 att/int region.
ACCESSION AY150027
VERSION AY150027.1 GI:28630421
KEYWORDS
SOURCE Bacteriophage pMLP1
ORGANISM Bacteriophage pMLP1
Viruses.

REFERENCE
AUTHORS Alexander,D.C., Devlin,D.J., Hewitt,D.D., Horan,A.C. and
Hosted,T.J.
TITLE Development of the Micromonospora carbonacea var. africana ATCC
JOURNAL 39149 bacteriophage pMLP1 integrase for site-specific integration
PUBMED in Micromonospora spp
12949170 Microbiology (Reading, Engl.) 149 (Pt 9), 2443-2453 (2003)
REFERENCE 2 (bases 1 to 2025)
AUTHORS Hosted,T.J. Jr., Alexander,D.C. and Hewitt,D.D.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2002) New Lead Discovery, Schering-Plough
Research, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA
FEATURES
source Location/Qualifiers
1..2025
/organism="Bacteriophage pMLP1"
/mol_type="genomic DNA"
/specific_host="Micromonospora sp. ATCC 39149"
/db_xref="taxon:219292"
/note="host is deposited in ATCC as Micromonospora
carbonacea var. africana Waltz et al."

```
misc_feature 1..2025
             /note="atc/int region"
CDS          197..622
             /code="xis"
             /codon_start=1
             /transl_table=11
             /product="excisionase"
             /protein_id="AA046044.1"
             /db_xref="GI:28630422"
             /translation="MNTPGLGRGTWAAVLTAREBAGITKSELARIQKDAITYGRW
EDGKRPDDADIVARVAQVIGLDLDELAAAGLRPGVTPTPATPTMDLDEIILVTRDP
KIDEMKRRIIALILERRERDRAAAIEETKRLLDLEFRS"
             628..1806
             /gene="intm"
             /gene_id="intm"
             /codon_start=1
             /transl_table=11
             /product="Integrase"
             /protein_id="AA046045.1"
             /db_xref="GI:28630423"
             /translation="MWIEKNQPVYRIIDVGRKVTIQTGYPTKSAXNMYQFRAEO
LOGNALMPRGQQTITADPFCGEMWPSYEKTKPTAVNSEGRINHLPLTGHITLDEL
DQGVITQWVNDLEAGVGPWPESTRGRKPLAKTISNGHGLHTTIGALIAARITBL
PCSTMLPRREPKEKELSDPEIGRLITALLPWRPLVMLVATGALWGEGALGRGR
VDLLARPLTVELOELASTGELVFOGSKTAKGRIVSFTTKVALILPLIAGKRS
DEVVFTAPRGAWRTNFRIMVACEEALPGLRHLDLHTHAALISAGRLSIS
RLGHSIAVTDLTGHIREVEGILAAIEEMAGVRADLEAEDELTDVLAADA
"
             join(1884..1895,1900..1911)
             1925..1986
             /note="atcp element"
             join(1969..1986,1991..2010)
ORIGIN
stem_loop   join(1884..1895,1900..1911)
misc_feature /note="atcp element"
stem_loop   join(1969..1986,1991..2010)

Query Match      50.6%; Score 125; DB 7; Length 2025;
Best Local Similarity 100.0%; Pred. No. 2,3e-14;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAGGGGAATCCACTCCGAGAGACGCCCGAGCAATCCGAGCANTAGCGAGCAACACGACG 60
DB 1825 TAGGGGAATCCACTCCGAGAGACGCCCGAGCAATCCGAGCANTAGCGAGCAACACGACG 1884

QY 61 GTCAAGTGCCTGTTGACCCCTTACCAAGGCGCCCGGTACGGGTTTCAATCCCATCGTC 120
DB 1885 GTCAAGTGCCTGTTGACCCCTTACCAAGGCGCCCGGTACGGGTTTCAATCCCATCGTC 1944

QY 121 ACCCG 125
DB 1945 ACCCG 1949

RESULT 5
LOCUS       AX338978                260 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION  Sequence 9 from Patent WO0187936.
ACCESSION  AX338978
VERSION    AX338978.1  GI:18129114
KEYWORDS   Micromonospora halophytica
SOURCE     Micromonospora halophytica
ORGANISM   Micromonospora halophytica
            Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
            Micromonosporinae; Micromonosporaceae; Micromonospora.

REFERENCE  1
AUTHORS   Hosted,T.J. and Horan,A.C.
TITLE     Isolation of Micromonospora carbonacea var africana pmlp1 integrase
and use of integrating function for site-specific integration into
Micromonospora halophytica and Micromonospora carbonacea chromosome
Patent: WO 0187936-A 9 22-NOV-2001;
JOURNAL   SCHERING CORPORATION (US)
FEATURES   Location/Qualifiers
            1..260
            /organism="Micromonospora halophytica"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:47864"
ORIGIN
Query Match      50.2%; Score 124; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 6,5e-14;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAGGGGAATCCACTCCGAGAGACGCCCGAGCAATCCGAGCANTAGCGAGCAACACGACG 145
DB 1 TAGGGGAATCCACTCCGAGAGACGCCCGAGCAATCCGAGCANTAGCGAGCAACACGACG 39149

QY 61 GTCAAGTGCCTGTTGACCCCTTACCAAGGCGCCCGGTACGGGTTTCAATCCCATCGTC 120
DB 61 GTCAAGTGCCTGTTGACCCCTTACCAAGGCGCCCGGTACGGGTTTCAATCCCATCGTC 124

QY 121 ACCC 124
DB 121 ACCC 124

RESULT 6
LOCUS       AY150028                145 bp      DNA      linear      BCT 02-S
DEFINITION  Micromonospora sp. ATCC 39149 atbB region.
ACCESSION  AY150028
VERSION    AY150028.1  GI:28630424
KEYWORDS   Micromonospora sp. ATCC 39149
SOURCE     Micromonospora sp. ATCC 39149
ORGANISM   Micromonospora sp. ATCC 39149
            Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
            Micromonosporinae; Micromonosporaceae; Micromonospora.

REFERENCE  1
AUTHORS   Alexander,D.C., Devlin,D.J., Hewitt,D.D., Horan,A.C. and
            Hosted,T.J.
TITLE     Development of the Micromonospora carbonacea var. africana A
39149 bacteriophage pMLP1 integrase for site-specific integr
in Micromonospora spp
Microbiology (Reading, Engl.) 149 (Pt 9), 2443-2453 (2003)
12949170
JOURNAL    PUBMED
PUBMED    2 (bases 1 to 145)
REFERENCE  2 (bases 1 to 145)
AUTHORS   Hosted,T.J., Jr., Hewitt,D.D. and Alexander,D.C.
TITLE     Direct Submission
JOURNAL    Submitted (12-SEP-2002) New Lead Discovery, Schering-Plough
Research, 2015 Galloping Hill Road, Kenilworth, NJ 07033, US
Location/Qualifiers
            1..145
            /organism="Micromonospora sp. ATCC 39149"
            /mol_type="genomic DNA"
            /strain="ATCC 39149"
            /db_xref="ATCC:39149"
            /db_xref="taxon:219305"
            /note="deposited in ATCC as Micromonospora carbonac
africana Waltz et al."
            1..145
            /note="atbB region"
            5..79
            /note="tRNA-His"
            56..80
            /note="attachment site; atbB/atp identity element"
            join(85..98,103..116)
ORIGIN
Query Match      36.4%; Score 90; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 2,3e-07;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 GGGTTCAATTCCTACGATGACCCGTCACGAGCCCGCTCCACTCGAGGGGCGCTCT 190
DB 56 GGGTTCAATTCCTACGATGACCCGTCACGAGCCCGCTCCACTCGAGGGGCGCTCT 190

QY 161 GCGTCTCTGAGGCTGCGGTCAGCGGT 190
DB 161 GCGTCTCTGAGGCTGCGGTCAGCGGT 190
```

Db 116 GGGCTTCCTGAGGTTGCGGTGACGGCGT 145

RESULT 7
LOCUS AY150032 143 bp DNA linear BCT 02-SEP-2003
DEFINITION Microcomospora nigra atcr region.
ACCESSION AY150032
VERSION AY150032.1 GI:28630428
KEYWORDS
SOURCE
ORGANISM Microcomospora nigra
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Microcomosporineae; Microcomosporaceae; Microcomospora.
REFERENCE
AUTHORS Alexander,D.C., Devlin,D.J., Hewitt,D.D., Horan,A.C. and
Hosted,T.J.
TITLE Development of the Microcomospora carbonacea var. africana ATCC
39149 bacteriophage PMLPI integrase for site-specific integration
in Microcomospora spp
JOURNAL Microbiology (Reading, Engl.) 149 (Pt 9), 2443-2453 (2003)
PUBMED 12949170
REFERENCE
AUTHORS 2 (bases 1 to 143)
Hosted,T.J., Hewitt,D.D. and Alexander,D.C.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2002) New Lead Discovery, Schering-Plough
Research, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA
LOCATION/Qualifiers
56.143
/organism="Microcomospora nigra"
/mol_type="genomic DNA"
/strain="ATCC 33088"
/db_xref="ATCC:33088"
/db_xref="taxon:145857"
/focus
1..55
/organism="Integration vector pSPRH840"
/mol_type="genomic DNA"
/db_xref="taxon:219291"
misc_feature
1..143
/note="atr region"
misc_recomb
55..56
/note="site of pSPRH840 recombination"
misc_feature
56..79
/note="pSPRH840 attachment site; atcr/atcrp identity
element"
stem_loop
join(94..105,110..121)

Query Match 32.0%; Score 79; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 CGGAGCAACGACGAGTGGCTGTGACCCCTGACAGGCGCCGGTACGGGTT 105
DB 1 CGGAGCAACGACGAGTGGCTGTGACCCCTGACAGGCGCCGGTACGGGTT 60
QY 106 CAATTCCCATAGTCACCC 124
DB 61 CAATTCCCATAGTCACCC 79

RESULT 8
LOCUS AY338477/c 33064 bp DNA linear HTG 07-OCT-2004
DEFINITION Streptomyces ambofaciens clone cosmid F6 strain ATCC 23877, ***
SEQUENCING IN PROGRESS ***; 2 ordered pieces.
ACCESSION AY338477 AY338478
VERSION AY338477.2 GI:53851167
KEYWORDS HTG; HTGS PHASE2.
SOURCE Streptomyces ambofaciens
ORGANISM Streptomyces ambofaciens
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE
AUTHORS Pang,X., Aigle,B., Girardet,J.M., Mangenet,S., Pernodet,J.L.,
Decaris,B. and Leblond,P.
TITLE Functional Angucycline-Like Antibiotic Gene Cluster in the Terminal
Inverted Repeats of the Streptomyces ambofaciens Linear Chromosome
JOURNAL Antimicrob. Agents Chemother. 48 (2), 575-588 (2004)
PUBMED 14742212
REFERENCE
AUTHORS 2 (bases 1 to 33064)
Aigle,B.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2003) Genetique et Microbiologie, UMR INRA 1128,
IFR 110, Universite Henri Poincare, Faculte des Sciences et
Techniques, Boulevard des Aiguillettes, BP 239,
Vandoeuvre-lès-Nancy 54 506, France
3 (bases 1 to 33064)
Aigle,B.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-2004) Genetique et Microbiologie, UMR INRA 1128,
IFR 110, Universite Henri Poincare, Faculte des Sciences et
Techniques, Boulevard des Aiguillettes, BP 239,
Vandoeuvre-lès-Nancy 54 506, France
Sequence update by submitter
On Oct 7, 2004 this sequence version replaced gi:39748104.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 13715: contig of 13715 bp in length
* 13716 13815: gap of unknown length
* 13816 33064: contig of 19249 bp in length.
LOCATION/Qualifiers
1..33064
/organism="Streptomyces ambofaciens"
/mol_type="genomic DNA"
/strain="ATCC 23877"
/db_xref="ATCC:23877"
/db_xref="taxon:1889"
/clone="cosmid F6"
239..1324
/gene="alpH"
239..1324
/gene="alpH"
/note="similar to Pfam PF00891 O-methyltransferase"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="putative O-methyltransferase"
/protein_id="AA30145.1"
/db_xref="GI:39748105"
/transl_table=1
/db_xref="GI:39748105"
RAVHVAVELVAALAEGLPASAELAGVADADALGRVRLAASNVFTRADGTE
LTPMADALRADHPSMSPRIALLMGHPHEDMSGFPERTVYTGSRALPKLGMARFEL
TKNARYGVFFRGMSMSASBETPTVAADVFTFTGVVDPCGQGGIAGILASSPAT
RGVLVPRVRENGAERFLAQGVADRTVAVGDLFVPPGADAVYVKKHIVHWPEBO
ALQILNNVRAIRKGRRLIAEMVIPLEOGDEPHSGKVLVDIMLVLVGRBRTPGQYAD
LIRAGFRLEERVETAAATSLVAAPV"
complement (2309..3835)
/gene="alpG"
complement (2309..3835)
/gene="alpG"
/note="similar to Pfam PF01494 FAD binding domain, and
PF01360 monooxygenase"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="putative oxygenase"
/protein_id="AA30146.1"

gene
CDS
/db_xref="GI:39748106"
/translation="MEGTADVDVIVVAGAGPTGLMAGELRGAARVVIETKLAAPTC
OSRGFTARAMEFDERGLPRFGDPLTASPMHFGAQDFVLEDAIFGAGI
PGQDTRAVIEDMAGELGADIRRGMEFVITLDGRODSVEITRTPOGQRTLRPL
VGCDDGSSSVRRPAGIDPFGTATQCMVTLADITGQVLTFRPLGERLNGMWAAPLSE
GVMRIIVCPDGRPAHDERKVTFOVAGAMHDTGGDISHGASWSSTTNATROAAH
YRGARVFLAGDAAHILPAGGQSLSTGVODANLGMKLAIVIRGDADTLDTYHAR
HVGARLLMNTAOGIVPLGGPESDPLRELMALVOYDVKRLAGIVSHLDIRYDLT
GDDTCDTGTGTHLVNGRMPPRTLVGAHGETTARLLHTGVLIDLADADAGIRRAA
GRPDVDAIVAHKPTDGPDLALGATAVLVPRPGVYAMAGTSFQGLEALQRMFGAR
"
/complement(3943..5418)
/gene="alpf"
/complement(3943..5418)
/gene="alpf"
/note="similar to Pfam PF01494 FAD binding domain, and
PF01360 monooxygenase"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative oxygenase"
/protein_id="AAR30147.1"
/db_xref="GI:39748107"
/translation="MTADHLSTDVVVVAGPVGMLAGELAHGVDVVLVEKHAAPT
ESRASTLHARTMEILDSRALIPERGDPNPRGHGGLPLDLTLPTSHPGQWKVQTR
TEVILQEWALSLGATLKCKHEVTLKOTGHEVEBAATGPOGTLRLHCRVLIACGSD
STVRLTGAPPGQDARBELRADVAGIDIPARRFERLETGLATAARRPDGVTVMH
BFGSTAGHRPGGDVSPBEITVAMKRVGEGDSSGTPMANAFGASROLTHYRGRLI
PAGRAHQMPIGGQALNTGVODAFNIGKALVVRGAPHTLIDTVHSEHGERV
LANIRAOGLMLGPPVGPVPRALALGELTLADVVRHLAGMTSGLDVRYDGGAPHL
GAPLPTCELRLRERTVTLARVRSQGVLLLTARGSPPARAGMDRVIEVAOFT
PGSALYTERVLRPDGHVAMAGSDGLTALTRWFGPPR"
/complement(5422..6366)
/gene="alpe"
/complement(5422..6366)
/gene="alpe"
/note="similar to Pfam PF03364 polyketide cyclase"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative bifunctional cyclase/dehydrase"
/protein_id="AAR30148.1"
/db_xref="GI:39748108"
/translation="MTREVEHEHTIGAPAAVVELADVTHMPRIFFPTIHVBLET
SGDERIRLWATANGOPKDMTSRRLIDARALTITRROIIPAPVHMGGTWIIIEBTA
DTSRRLIDVYALGDDPHDLMTBOAVKSTOELAAKNAVEAAHAAATELPESE
TDTVILAGAKDVDFINEANHLMSERLPHVAVRLTEDIPGQLEMDTRAODGSVHT
TKSVYVFPHHKIAVKOVLTPALMTLHTEWTFTEDEOQTSSOHTVTLNTANLARI
LQDATTADARAIVHSAIYSTNSATLAHKAHAQKKA"
/complement(6419..7207)
/gene="alpd"
/complement(6419..7207)
/gene="alpd"
/note="similar to Pfam PF00106 short chain dehydrogenase"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative polyketide ketoreductase"
/protein_id="AAR30149.1"
/db_xref="GI:39748109"
/translation="MTDTTQRAVAVITGATSGIGLASARLLGAOGRVFIARSADV
AATVKELOGEGLDVGDVAVVDVDTASVDAMIQAADVRFGLDLYVNNAGSGGPTPD
IADELMNDVIDTINLSVFRALATLIGKLRANDRGITINASTAGQGVVLAAPISA
SKHGIVGFTALGNELAPGTITVAVCPGVETPMARVQGAATAADTSEALIDKE
QAKIPLGRYSTPOVAGLIGVIASTDAAITSQALNVCGLGNF"
/complement(7220..7489)
/gene="alpc"
/complement(7220..7489)
/gene="alpc"
/note="similar to Pfam PF00550 phosphopantetheine
attachment site"
/codon_start=1
/evidence=not_experimental

gene
CDS
/transl_table=11
/product="putative acyl carrier protein"
/protein_id="AAR30150.1"
/db_xref="GI:39748110"
/translation="NATLFTDDEKRLRDAAVAGEVDLIDTEFEF
ALLEGSLIEREYGLSDBEANGENTRTFIEVNAQALPAQAA"
/complement(7558..8769)
/gene="alpb"
/complement(7558..8769)
/gene="alpb"
/note="similar to Pfam PF00109 and PF02801 beta-keto-
synthase"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative chain length determinant"
/protein_id="AAR30151.1"
/db_xref="GI:39748111"
/translation="MTASVVTGIGVTAENGELKDYMAATLGGKIGIRI
YPARLAGQIDFERADRLPSRLPOTDRVTRLALVADNALADAGADPAOL
ITASAGFERFGGEGLOALMSOGSOVVAVSFAMVAVNSGOISIRMGK
SEAGGLDAVQAQRQIRKGTSLITGANDASICPMGVAQLAGRLSTD
PFPRDARGFVPGGAILAEDAAARTGARPYGETIAGATIDPRPGSG
AIFETALADARLSAIDVIDVADGADGAPGIRADALSTVFGRGVPTVP
YSGAGPDLIAAFLALRDVTPPTVHIDPCADYPLDVLGEPAPRLTAL
GFNSAAVRAV"
/complement(8766..10037)
/gene="alpa"
/complement(8766..10037)
/gene="alpa"
/note="similar to Pfam PF00109 and PF02801 beta-keto-
synthase"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative ketosyl synthase"
/protein_id="AAR30152.1"
/db_xref="GI:39748112"
/translation="MSRRVVTIGIEVAPGVGRENFWLLSGRTATR
PFRSVAEADFPDPAAGLSPOEVRILDRAPQFVAVSRGAVADSGDLVAA
VTGSAVGAITMGLDBEYRVVDDGRLVADHTVAVPHDYDVMSFPAE
GPTSVSTGCTSGIDSVIAYELVREGSADVIVGSSAPISPTMACFSD
HDEPCASRPDPKTRNGVILGEGCAFVLEEDLSARKGAHIAETIACS
TGIRDGAEMEAIDLALAEARLNDQALIDVNAHSGSKQNDREHAAFK
RTPVSIKSMWGHSLGALGSIEMASALAMEYDVVPPANLHTADPECL

Query Match 17.4%; Score 43; DB 2; Length 33064;
Best Local Similarity 50.7%; Pred. No. 44;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps

QY 44 GACGAGCAACCAAGAGTCAAGTGGCTCTTGAACCCCTGACAGAGGCCCGGATAC
DB 21095 GCGTGGCGCCGACCTGCGGACGCGCTCAAGCCCAAGGCGCAATCCCAACAC
QY 104 TTCAATTCCTCAGTCAAGTCAAGTGGCTCTTGAACCCCTGACAGAGGCCCGGATAC
DB 21035 TACCGGCTTTCGCTGCGGACGCGGACGCGCAACCTGCGGAGGCGCGGCTTTC
QY 164 GTTCTGAGAGGTTTCGCGTCAAGTGGCTCTTGAACCCCTGACAGAGGCCCGGATAC
DB 20975 GTACTGAGAGGCGCGGCTTTCGCGTCAAGTGGCTCTTGAACCCCTGACAGAGGCGCGGCTTTC
QY 224 GGGAGTGGCTTCGCGCTTCGCGG 246
DB 20915 GCGTGGCGCTTCGCGCTTCGCGG 20893

RESULT 9
AC079680/c AC079680 175217 bp DNA linear ROD 15
LOCUS M8 musculus chromosome 10 clone rp23-282p18 strain C57BL/6
DEFINITION complete sequence.
ACCESSION AC079680

VERSION AC079680.34 GI:25013348
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 175217)
 AUTHORS Yang, L., Hu, P., Wu, H., Prescott, A. and Roe, B.A.
 TITLE Mus musculus BAC Clone rp23-282p18
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 175217)
 AUTHORS Yang, L., Hu, P., Wu, H., Prescott, A. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (07-SEP-2000) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 3 (bases 1 to 175217)
 AUTHORS Yang, L., Hu, P., Wu, H., Prescott, A. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAY-2002) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 4 (bases 1 to 175217)
 AUTHORS Yang, L., Hu, P., Wu, H., Prescott, A. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAY-2002) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 5 (bases 1 to 175217)
 AUTHORS Yang, L., Hu, P., Wu, H., Prescott, A. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (15-NOV-2002) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 COMMENT On Nov 15, 2002 this sequence version replaced gi:21070690.
 ----- Genome Center
 Center: Department of Chemistry And Biochemistry
 The University of Oklahoma
 Center code:UOKNOR

 FEATURES
 source
 1.175217
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /chromosome="10"
 /clone="rp23-282p18"
 /clone_1b="RP23 - 23 Female (C57BL/6J) Mouse BAC Library"
 /note="This is one of two clones in the same well from rp23-282p18"

 ORIGIN
 Query Match 17.2%; Score 42.6; DB 10; Length 175217;
 Best Local Similarity 61.1%; Pared. No. 32;
 Matches 69; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

 QY 134 GGGCCCCCTCACTCGAGGGGCGCTTCGCGGCTTCGAGGCTTCGCGGCGCGCTCGG 193
 Db 100502 GGGCTGTGTCGCTCGAGAGGGCGGCTGCGGGGCGGGGCGCGGCGCGCATCGG 100443
 QY 194 CTCGGCGCTCGGGGAGCTCGGCCCGCTCGCGGAGTGGCTTCGCGGCTCGCGGG 246
 Db 100442 CGCGCGGGGCGGGGCGGGGCGCGCGCGGAGGCGGAGTGGCGAGGCTTCGCGG 100390

 RESULT 10
 LOCUS AC116464 194452 bp DNA linear HTG 09-MAR-2004
 DEFINITION Mus musculus chromosome 10 clone RP23-60H16 map 10, WORKING DRAFT
 AC116464
 AC116464.6 GI:4486394
 VERSION

KEYWORDS HTG, HTGS PHASE2, HTGS DRAFT, HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 194452)
 AUTHORS Birren, B., Nussbaum, C. and Lander, E.
 TITLE Mus musculus chromosome 10, clone RP23-60H16
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 194452)
 AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fardo, S., Ford, S., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 194452)
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Fardo, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Liu, X., Manning, J., Matthews, C., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spence, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Mar 2, 2004 this sequence version replaced gi:24182084.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence submissions@genome.wi.mit.edu

 Project Information
 Center project name: l23285
 Center clone name: 60 H 16

 Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.

* in its sequence will be deposited as soon as it is available and
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 25084: contig of 25084 bp in length
 * 25085 25184: gap of 100 bp
 * 25185 37881: contig of 12637 bp in length
 * 37822 37921: gap of 100 bp
 * 37922 120136: contig of 8215 bp in length
 * 120137 120236: gap of 100 bp
 * 120237 128630: contig of 8394 bp in length
 * 128631 128730: gap of 100 bp
 * 128731 174500: contig of 45770 bp in length
 * 174501 174600: gap of 100 bp
 * 174601 194432: contig of 19852 bp in length.

```

1. .194452
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10"

```

```
misc_feature
    /clone="RP23-60H16"
    /clone_lib="RPC1-23 Female Mouse BAC
    1.25084
    /data="assembly fragment
```

```
25185. .37821
/note="assembly_fragment
37922 120136
```

```
misc_feature 37922..120136 :  
/note="assembly_fragment  
120237..128630  
/date="20060601" fragment
```

```
misc_feature 128731. 174500
misc_feature /note="assembly_fragment
174601. 194452
misc_feature "assembly_fragment
```

```

/notes="assembly_fragment
clone_end:T7
vector_side:right"

```

17.28; Score 42.6; DB 2; Length 194452,

Best Local Similarity	61.1%;	Pred. No. 32;	
Matches	69;	Conservative	0;
		Mismatches	44;
		Indels	0;
		Gaps	0

QY 134 GACCCCTTCACATCGAGAGGAGGAGCTTCGAGCTTCTTAGAGCTTCCGAGGTCGAGCGATCCG 193
Db 176261 GGCTCTGTCGCGTCGAGAGCGGAGCTTCGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCGAGCGATCCG 176320

Qy 194 CTGGGCGCTGGAGGAGCTCGGCGCCGGTCGGCGAGAGTGGGCTCTGGCGCTTCGGGG 216

Ddb 176321 CGGCGCGGGGCGGGGGCGGGCGCCGAGGCGGGGATGGCGAGGGCTTTCGGCG 176373

RESULT 11		
SEABCT/c		
LOCUS	1976 bp	DNA
SEABCT		linear
		BCT 07-DEC-1995

REFERENCE
O'Neill, M.P., Bady, E.A., Radford, A., Baumberg, S. and Cove, J.
(1995) The use of PCR to isolate a putative ABC transporter from
Saccharopolyspora erythraea
FEMS Microbiol. Lett. 131 (2), 189-195 (1995)

MEDLINE	96004570
PUBMED	7557329
REFERENCE	2 (bases 1 to 1976)
NUMBERS	O'Neill, M.

TITLE Direct Submission
JOURNAL Submitted (29-VII-1994) M. O'Neill, Department of Microbiology
University of Leeds Leeds LS2 9JT, UK

source	FEATURES
1. . . 1976	Location/Vuallititels
/swatzi sm-"Saccharono]vsnora erythraea"	

-10_signal

RBS

gene

CDS

1

OPTCTN

Query

Matche

Qy

Db

QY

55

PESTHIT

1

1

LOCUS AF315943 36654 bp DNA linear PRI 02-JAN-2001
 DEFINITION Homo sapiens WNT6 (WNT6) gene, partial cds; and WNT10A (WNT10A) gene, complete cds, complete sequence.
 ACCESSION AF315943
 VERSION AF315943.1 GI:12007357
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE AUTHORS Rump, A., Hayes, C., Brown, S.D.M. and Rosenthal, A.
 TITLE 1 (bases 1 to 36654)
 GENOMIC Rump, A., Hayes, C., Brown, S.D.M. and Rosenthal, A.
 SEQUENCE Genomic sequence of the Wnt6 gene and the Wnt10a gene from human
 Q35
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 36654)
 AUTHORS Rump, A., Hayes, C., Brown, S.D.M. and Rosenthal, A.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-2000) Genome Analysis, Institute of molecular
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
 FEATURES
 SOURCE
 1. 36654
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q35"
 /clone="cosmid 63C20"
 <1. 7890
 /gene="WNT6"
 /note="Intron/exon structure determined by alignment of
 human genomic sequence with murine wnt6 mRNA sequence
 presented in Genbank Accession Number M89800"
 join(<4700..4920,5158..5492,7057..7890)
 /gene="WNT6"
 /product="WNT6"
 join(<4700..4920,5158..5492,7057..7518)
 /gene="WNT6"
 /codon_start=2
 /product="WNT6"
 /protein_id="AA05154.1"
 /db_xref="GI:12007359"
 /translation="AVGSLPMDPTSCIKARLRLAGQAEICQAEVVAELARGARL
 LGVECOFQPRFRWNCSSHSKAFRILODIRAFATATAGASNAVTAQASMGEL
 LQCCOAPRGRRPBGSLGTPGPPRPPASPPGSAWEMGGGDDVDDEKSRLEM
 DARRKRGRIALVQLAHNNAERLAVSRHTRECKCHGSGCALRTCKWCKPPRE
 VGARLIERFHGASRWMTDNGKALLPVRLTKPRADILYADSDPCAPNRTGSP
 GTRGRACNSAPDLSCDILCCGRGRHRSVLEENCLCRFHWCVVQCHRCVKREL
 SLCL"
 14400..27595
 /gene="WNT10A"
 /note="Intron/exon structure determined by alignment of
 human genomic sequence with murine wnt10a mRNA sequence
 presented in Genbank Accession Number U61969"
 join(14440..14777,15830..16092,23653..24032,26436..27595)
 /gene="WNT10A"
 /product="WNT10A"
 join(14665..14777,15830..16092,23653..24032,26436..26933)
 /gene="WNT10A"
 /codon_start=1
 /product="WNT10A"
 /protein_id="AA045153.1"
 /db_xref="GI:12007358"
 /translation="MGSAPRPRLRLRPOPAPALWLPFLITLAAAMPRSANDI
 LDRLEPPLVNAANTVCLTLPGLSRQMEVCVRHPDVAAAGIAGIQTITHEQFRD
 ORWNCSELETRNKIPYESPPIRGRBSAFATIAAGVAVASNAALGKLGKAGCD
 ASRGDEARPRKRLRLDIALQKGLSHGVPEPALPTASPGLDGEMWGGSPDM
 GGFGRFSDPLDSRPHRDIAHRMLNNRVQAVMMNRKCKCHTSGSCOLKTC
 WQVPEFTVALALSRFHATLIRPHNRNGOLEPGAGAPSPAPGAPGRRRASP
 DLVFEKSPDCEPRLDSAGTVRLCNKSASGSDGSCNCCGRGHNTITQTSERC
 HCRFHWCCFVCEBCRTIEWSVCK"

Query Match 16.8%; Score 41.4; DB 9; Length 36654;
 Best Local Similarity 54.2%; Pred. No. 86;
 Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 77 ACCCCTGACCGAGGCCCGGTGACCGGTTCAATTCCATCAGTACCCGTACAGGAAGC 136
 Db 20447 AGCTCTGTTCAAGGGCCAGATTTCGTTCTGCGCCACGCTCTGAGGCAACACAGCC 20388
 QY 137 CCCCTCACTGGAGAGGGCCCTTCGCGCTTCTGAGGGTTGCGGTCAGCGGTCGCTC 196
 Db 20387 ACTGTCCATCTGGCCAGGTGAGAGGGGAGCGCGGTGGGCGCAATGCGAGCTGGCGTCC 20328
 QY 197 GGCGCTGGGAGACCTGCGCCCGCTGCGGAGAGTGG 231
 Db 20327 GGAGCAGGAGGACGCGGTGGGGGGTGGTGGGG 20293
 RESULT 13
 AC073128/c
 LOCUS AC073128 55590 bp DNA linear PRI 14-JUL-2002
 DEFINITION Homo sapiens BAC clone Rpl1-64705 from 2, complete sequence.
 ACCESSION AC073128
 VERSION AC073128.5 GI:21694010
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE AUTHORS Sulston, J.E. and Waterston, R.
 TITLE 1 (bases 1 to 55590)
 JOURNAL Toward a complete human genome sequence
 MEDLINE Genome Res. 8 (11), 1097-1108 (1998)
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 55590)
 AUTHORS Harris, A. and Kozlowski, A.
 TITLE The sequence of Homo sapiens BAC clone Rpl1-64705
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 55590)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-JUN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 55590)
 REFERENCE 4 (bases 1 to 55590)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (14-MAY-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 55590)
 REFERENCE 5 (bases 1 to 55590)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 6 (bases 1 to 55590)
 REFERENCE 6 (bases 1 to 55590)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jul 4, 2002 this sequence version replaced gi:20564458.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: saplens@wustl.edu

 Summary Statistics
 Center project name: H_NH0647005

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatem, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-459119, 2000 bp overlap; the clone sequenced to the right is RP11-3304, 2000 bp overlap. Actual start of this clone is at base position 74819 of RP11-459119; actual end is at base position 13291 of RP11-3304.

Data from AC092152 was used to finish this clone, AC092152.

FEATURES

SOURCE

Location/Qualifiers
1. 55590
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-64705"
/clone_lib="RP11-11"
/clone_lib="RP11-11"
2525. 2634
/rpt_family="CR1"
3166. 3478
/rpt_family="Alu"
3581. 3680
/rpt_family="MIR"
4816. 4888
/rpt_family="GC_rich"
7951. 8074
/rpt_family="MIR"
8075. 8404
/rpt_family="MaLR"
8405. 8717
/rpt_family="Alu"
8718. 8931
/rpt_family="MaLR"
8932. 9040
/rpt_family="MIR"
9074. 9526
/rpt_family="MaLR"
10409. 10473
/rpt_family="MIR"
11135. 11168
/rpt_family="(GCCTG)n"
11338. 11394
/rpt_family="(CTAGGG)n"
11764. 11917
/rpt_family="MER1_type"
12603. 12642
/rpt_family="MIR"

repeat_region 12957. 13091
/rpt_family="MIR"
misc_feature 13493. 14286
/note="CpG island (%GC=73.4, o/e=0.80, #CpGs=75)"
repeat_region 14769. 14797
/rpt_family="(TG)n"
repeat_region 16760. 16762
/rpt_family="(CCCCG)n"
repeat_region 16939. 16999
/rpt_family="(CCCCG)n"
repeat_region 17113. 17157
/rpt_family="GC_rich"
repeat_region 19095. 19336
/rpt_family="MIR"
repeat_region 19688. 19742
/rpt_family="C-rich"
repeat_region 20890. 21012
/rpt_family="CR1"
repeat_region 21007. 21255
/rpt_family="MIR"
repeat_region 21262. 21330
/rpt_family="T-rich"
repeat_region 22145. 22167
/rpt_family="(TTTA)n"
repeat_region 22168. 22771
/rpt_family="L1"
repeat_region 23039. 23181
/rpt_family="MIR"
repeat_region 23727. 23807
/rpt_family="L2"
repeat_region 23935. 24014
/rpt_family="MIR"
repeat_region 24250. 24472
/rpt_family="MIR"
repeat_region 25532. 25588
/rpt_family="(CA)n"
repeat_region 26178. 26248
/rpt_family="GC_rich"
repeat_region 26599. 26656
/rpt_family="CR1"
repeat_region 26970. 27723
/rpt_family="L2"
repeat_region 30381. 30423
/rpt_family="GC_rich"
repeat_region 31952. 32271
/rpt_family="L1"
repeat_region 32404. 32698
/rpt_family="Alu"
repeat_region 33184. 33244
/rpt_family="L1"
repeat_region 33248. 33301
/rpt_family="L1"
repeat_region 33302. 33647
/rpt_family="Alu"
repeat_region 33648. 34058
/rpt_family="L1"
repeat_region 34048. 34439
/rpt_family="L1"
repeat_region 34440. 34522
/rpt_family="(CTA)n"
repeat_region 34523. 34624
/rpt_family="L1"
repeat_region 34879. 35185
/rpt_family="Alu"
repeat_region 35196. 35571
/rpt_family="L2"
repeat_region 35578. 35682
/rpt_family="Alu"
repeat_region 35691. 35951
/rpt_family="L2"
repeat_region 35955. 36108
/rpt_family="Alu"
repeat_region 36132. 36381

Query Match 16.8%; Score 41.4; DB 9; Length 55590;
 Best Local Similarity 54.28; Pred. No. 77;
 Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 77 ACCCCCTGACACGAGGCGCCCGTACGGGTTCCATGTCATGACACCCGACGAGGC 136
 DB 20006 AGCTCTGTTACAGGCGCCCGATTCGTTGCGCCCGACGTCGTGCGGACACAGGC 19947
 QY 137 CCCCTCACTCGAGAGGCGGCGCTTCGCGCTTCGAGGCTTCGAGCGCGTGGCTC 196
 DB 19946 ACTGTCCCATCTGCGCAGCTGAGAGGGAGCGGCTGGGCGCAGCGCAGTGGCTGCC 19887
 QY 197 GGGCGTGGGAGACTCGCGCCCGTTCGCGCGGAGTGG 231
 DB 19886 GAGAGGAGGAGCGCGTGGCGGCGTGGTGGTGGG 19852

RESULT 14

AC027416 152129 bp DNA linear HTG 07-JUN-2000
 LOCUS AC027416.2 GI:8317289
 DEFINITION Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
 UNORDERED PIECES.

AC027416
 HTG: HTGS_PHASE1, HTGS_DRAFT.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE 1 (bases 1 to 152129)
 JOURNAL Biren, B., Linton, L., Nusbaum, C. and Lander, E.
 REFERENCE Homo sapiens, clone RP11-504G11
 AUTHORS Unpublished

2 (bases 1 to 152129)
 Biren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barn, N., Bastien, V., Bedalov, F.,
 Boguslavsky, L., Bouckhagalter, B., Brown, A., Burkett, G.,
 Campoliano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dextrallano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Donno, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kahn, L., Karatas, A.,
 Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marcis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPherson, R.,
 Meldrum, J., Menus, L., Milnova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisanu, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zaitoun, J., Zimmer, A. and Zody, M.
 Direct Submission

TITLE
 JOURNAL
 COMMENT
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L7458
 Center clone name: 504_G 11
 ----- Summary Statistics -----
 Sequencing vector: M13; M7815, 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 135376 bases at least Q40
 Consensus quality: 143264 bases at least Q30
 Consensus quality: 146503 bases at least Q20
 Insert size: 161000; agarose-fp
 Insert size: 149029; sum-of-contigs
 Quality coverage: 3.1 in Q20 bases; agarose-fp
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 32 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1	1005:	contig of 1005 bp	in length
1006	1105:	gap of 100 bp	
1106	2402:	contig of 1297 bp	in length
2403	2502:	gap of 100 bp	
2503	3823:	contig of 1321 bp	in length
3824	3923:	gap of 100 bp	
3924	5020:	contig of 1097 bp	in length
5021	5120:	gap of 100 bp	
5121	6161:	contig of 1041 bp	in length
6162	6261:	gap of 100 bp	
6262	7547:	contig of 1286 bp	in length
7548	7647:	gap of 100 bp	
7648	9983:	contig of 2336 bp	in length
9984	10083:	gap of 100 bp	
10084	12556:	contig of 2473 bp	in length
12557	12656:	gap of 100 bp	
12657	15043:	contig of 2387 bp	in length
15044	15143:	gap of 100 bp	
15144	17123:	contig of 1980 bp	in length
17124	17223:	gap of 100 bp	
17224	19466:	contig of 2243 bp	in length
19467	19566:	gap of 100 bp	
19567	21928:	contig of 2362 bp	in length
21929	22028:	gap of 100 bp	
22029	24319:	contig of 2291 bp	in length
24320	24419:	gap of 100 bp	
24420	27059:	contig of 2640 bp	in length
27060	27159:	gap of 100 bp	
27160	30170:	contig of 3011 bp	in length
30171	30270:	gap of 100 bp	
30271	33968:	contig of 3698 bp	in length
33969	34068:	gap of 100 bp	
34069	38179:	contig of 4111 bp	in length
38180	38279:	gap of 100 bp	
38280	42366:	contig of 4087 bp	in length
42367	42466:	gap of 100 bp	
42467	46365:	contig of 3899 bp	in length
46366	46465:	gap of 100 bp	
46466	51285:	contig of 4820 bp	in length
51286	51385:	gap of 100 bp	
51386	55871:	contig of 4486 bp	in length
55872	55971:	gap of 100 bp	
55972	60595:	contig of 4624 bp	in length
60596	60695:	gap of 100 bp	
60696	66595:	contig of 5900 bp	in length
66596	66695:	gap of 100 bp	
66696	73218:	contig of 6523 bp	in length
73219	73318:	gap of 100 bp	
73319	77115:	contig of 3797 bp	in length
77116	77215:	gap of 100 bp	
77216	85022:	contig of 7807 bp	in length
85023	85122:	gap of 100 bp	
85123	93414:	contig of 8192 bp	in length
93415	93414:	gap of 100 bp	
101193:	contig of 7779 bp	in length	

[illegible]

/transl_table=11
/product="hypothetical protein SCC117.03."

Query Match 16.8%; Score 41.4; DB 1; Length 310550; *
Best Local Similarity 58.5%; Pred. No. 47; Mismatches 51; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 124 CCTACACGAAAGGCCCTCCACTCGGAGGGGCGCTTCGGCGTTCCTGAGGGTTCCGGTC 183
DB 53337 CGACTTGGAGGCGGCTCGCGCTACCGCGACGGCTCGGTGCGGGTGGC 53396
QY 184 AGGCGGTCGGCTCGGCGCTGGGGAAGTCCCGTCCGGCGGAGTGGCTCGGCGTCCG 243
DB 53397 CGCGGTCACGCGTCCGGGTCCCTCGGCGCAGAGCGCGGCGTGAACCTTCTCGACCA 53456
QY 244 GGG 246
DB 53457 GCG 53459

Search completed: April 15, 2005, 16:46:06
Job time: 1940 secs

THIS PAGE BLANK (USPTO)

Mon Apr 18 09:47:34 2005

us-09-855-340b-6.rst

Pa

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 15:45:55 ; Search time 3283 Seconds
(without alignments)
2863.806 Million cell updates/sec

Title: US-09-855-340B-6
Perfect score: 247
Sequence: 1 tagggggaatccatccgagag.....gtggcctcgagcgcg999a 247

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb est1:*
2: gb est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.2	19.9	925	CNS0091P	AL053013 Drosophila
2	46	18.6	457	AV432456	AV432456 AV432456
3	45.8	18.5	932	CNS0072Q	AL066742 Drosophila
4	44	17.8	1229	BMS54300	BMS54300 AGENCOURT
5	43.2	17.5	1073	B0721449	B0721449 AGENCOURT
6	42.2	17.1	958	AG056804	AG056804 Pan. trogl
7	42.2	17.1	961	CNS01087	AL098641 Drosophila
8	42	17.0	934	CF753549	CF753549 EST-Conti
9	41.8	16.9	865	AG136795	AG136795 Pan. trogl
10	41.2	16.7	726	CL147646	CL147646 104 327.1
11	41.2	16.7	826	CG270532	CG270532 OXKAE40TH
12	41.2	16.7	839	CC626818	CC626818 OGPARE2TH
13	41.2	16.7	925	CNS0091P	AL053013 Drosophila
14	41	16.6	738	BI889087	BI889087 ZF637-2-0
15	40.4	16.4	345	CE451542	CE451542 l19t-gss-
16	40.4	16.4	1288	BZ579414	BZ579414 msh2 6344
17	40	16.2	754	CC658523	CC658523 OGVF68TH
18	39.8	16.1	814	CB313315	CB313315 AGENCOURT
19	39.8	16.1	935	CNS006XK	AL066051 Drosophila
20	39.8	16.1	1312	AG080972	AG080972 Pan. trogl
21	39.6	16.0	598	BE456076	BE456076 HVSMEG001
22	39.6	16.0	977	CNS00JX7	AL076850 Drosophila
23	39.4	16.0	914	BU595137	BU595137 AGENCOURT
24	39.4	16.0	963	CK424745	CK424745 AUF_1psto

25	39.4	16.0	1051	9	CL482997
26	39.2	15.9	391	8	BH217504
27	39.2	15.9	844	9	CNS0052P
28	39.2	15.9	925	8	B09283
29	39	15.8	773	7	CN016299
30	39	15.8	1367	7	CV068980
31	38.8	15.7	870	9	CC626829
32	38.6	15.6	447	2	BF145773
33	38.6	15.6	581	6	CB659032
34	38.6	15.6	696	9	CL967378
35	38.6	15.6	1115	5	BM24406
36	38.4	15.5	305	1	AT193646
37	38.4	15.5	376	5	BU996209
38	38.4	15.5	405	5	BU999117
39	38.4	15.5	434	5	BU997816
40	38.4	15.5	498	6	CB859042
41	38.4	15.5	522	7	CF659480
42	38.4	15.5	602	6	CB875859
43	38.4	15.5	616	6	CA031301
44	38.4	15.5	646	6	CD422925
45	38.4	15.5	650	7	CN149220

ALIGNMENTS

RESULT 1
CNS0091P/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TERT end of BACR19D16 of RPL1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL053013
GI:4934461
VERSION
GSS.
KEYWORDS
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequen
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope

COMMENT

- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (funded by the Howard Hughes Medical Institute) and the Drosophila melanogaster genome using these BACs. For further information, please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mamooser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPL1-98 and was constructed by partial digestion of Drosophila DNA provided by the BDGP from isogenic strain Y2; on bw sp, the same strain used for the pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, filters for hybridization from the BACPAC Resource Center is found at http://bacpac.med.buffalo.edu/drosophila_bac.htm

FEATURES

source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_11b="RPL1-98"
/note="end : TERT"

ORIGIN

Query Match 19.9%; Score 49.2; DB 9; Length 925;
Best local similarity 12.0%; Pred. No. 0.043;

Query Match	17.0%;	Score 42;	DB 7;	Length 934;
Best Local Similarity	64.3%;	Pred. No. 2.8;		
Matches 63;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0

QY 148 GGAGGGGGCTTGGCGCTTCTGAGGGTTCGGCGTCAAGCGGCTCGGCTCGCGCTGCGGG 207
 DB 660 GGAGGGAGCGCGCGGCTTCTTGGCGCGCGGGAGAGGGGGCGCGCGCGCGGG 719
 QY 208 ACTCGCCCCCTTCGGCGGAGTGTGCTTCGCGCTCGCGG 245
 DB 720 GGGCGACCCCGCGCGCGGGGGCGGGGAGATCGGG 757

RESULT 9

AG136795 865 bp DNA linear GSS 04-NOV-2001
 LOCUS Pan troglodytes DNA, clone: PTB-150H18.F, genomic survey sequence.
 DEFINITION AG136795
 ACCESSION AG136795
 VERSION AG136795.1 GI:1666473
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of library PTB
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 865)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Ageo Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenhiro-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.
 PRIMERS
 Sequencing: -21M13
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
 Location/Qualifiers
 1. 865
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-150H18.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES

ORIGIN
 Query Match 16.9%; Score 41.8; DB 9; Length 865;
 Best Local Similarity 63.4%; Pred. No. 3.1;
 Matches 64; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 146 TCGAGGGGGCTTGGCGGCTTCTGAGGGTTCGGCGTCAAGCGGCTCGGCTCGCGCTGGG 205
 DB 542 TCGGGGGGGCGCTGCGCGCGGCTGCGCGCGCGCGCGCGCGCTGCGCGCGCGGGG 483
 QY 206 GGAATCGGGCGGCTGCGCGGAGTGGCTTCGCGCTCGCGGG 246
 DB 482 GGGCGTGGCGCGCGCGCGCGCTGCGCTTGGGGGCGCGGG 442

RESULT 10
 LOCUS CL147646 726 bp DNA linear GSS 06-JAN-2004
 DEFINITION 104 327 10592533.116.31781.041 Sorghum methylation-filtered library
 (LibID: 104) Sorghum bicolor genomic clone 10592533, genomic survey

ACCESSION

sequence.
 CL147646
 CL147646.1 GI:40647662
 GSS.
 Sorghum bicolor (sorghum)
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Paniceae;
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

1 (bases 1 to 726)
 Budiman, M.A., Flick, E., Jones, J., Nunberg, A., Citek, R.W.,
 Robbins, D., Rohlfing, T., Bradford, K., Fries, J., McWenamy, J.,
 Trani, L., Isak, A., Zimmerman, C., Lakey, N. and Bedell, J.A.
 Genethresher methylation filtered genomic sequences from Sorghum
 bicolor
 Unpublished (2004).
 Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: 327 row: b column: 17
 Seq primer: T3 Reverse
 Class: shotgun
 High quality sequence stop: 726.

JOURNAL

Unpublished (2004).
 Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: 327 row: b column: 17
 Seq primer: T3 Reverse
 Class: shotgun
 High quality sequence stop: 726.

FEATURES

source

1. 726
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /cultivar="ATx623"
 /db_xref="taxon:4558"
 /clone="10592533"
 /clone_lib="Sorghum methylation-filtered library (L104)"
 /note="Organ: leaf; Vector: pBSK(-); Site: 1: HincI
 prepared from purified nuclei was randomly sheared,
 end-repaired, size fractionated to enrich for the 0
 kb fraction, ligated into HincI-digested pBSK(-)
 and electroporated into E. coli cells. This is a
 methylation-filtered library."

ORIGIN

Query Match 16.7%; Score 41.2; DB 9; Length 726;
 Best Local Similarity 49.5%; Pred. No. 4.4;
 Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 31 CAATCCGAGCATACGAGCAACACAGGTCAGTGGCTGTGACCCCTGACCA
 DB 393 CCAGCGCGGCTACGCGCGGAGCGGCTCCGACGCGGGTAGTCCCGGACAA
 QY 91 GCCCGGATGAGGGTTCAATCCATCAGTCAACCGTACACAGGCCCTTCACCTC
 DB 453 CATTAAGCGGAGACAGACGCGCATCAAGCGCTCCGAAACCGGCGCGCTCGCGCTC
 QY 151 GGGGCGCTTGGCGGCTTCTGAGGGTTCGCGGTCAAGCGGCTCGGCGCTCGGGG
 DB 513 CGGCGGAGCGGCTTCACGCTTCCTCCCGCGGAGCAGCGCTCGAGCTCGTGG
 QY 211 CGGCGCGTTCGCGGAGATGCGCTTCGCGCTCGG 244
 DB 573 CGGCGTGGCGCGCGGACGCGCGCGGAGTGGCGG 606

RESULT 11

CG270532 826 bp DNA linear GSS 25-JAN-2004
 LOCUS CG270532 826 bp DNA linear GSS 25-JAN-2004
 DEFINITION CG270532 ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA058560
 genomic survey sequence.
 ACCESSION CG270532
 VERSION CG270532.1 GI:34182673
 KEYWORDS GSS.
 SOURCE Zea mays

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 826)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OXAE40TV
Contact: Cathy Whitelaw

FEATURES
source
1. 826
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMEMA0585607"
/clone_1ib="ZM_0.7.1.5_KB"
/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN
Query Match 16.7%; Score 41.2; DB 9; Length 826;
Best Local Similarity 49.5%; Pred. No. 4.4;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 31 CAATCCGAGCATGACGAGCAACAGCAGGTGAGTGGCTGTGACCCCTGACACAG 90
DB 248 CGACGCGCGGCTACGCGCGCGCGCGCTCCGACGCGCGGTAGCGCCCGACAG 307
QY 91 GCGCCGATGAGGTTCAATTCCTCATGTCACCCGTACAGAGCCCTCCACTCGGA 150
DB 308 CGTCCAGGCGGAGCAGCAGCGGCATGAGCGCTCGGAGAACCGGCGCGCTCGT 367
QY 151 GGGGGCTTGGCGCTTCTGAGGGTTCCGCGTCAAGCGCGCTCGCGCTGGGGAGCT 210
DB 368 CGCCGCGGACCGCTCCACGCTCTCTGCTCCGCGGAGAGCGCTCCAGCGCGCTGCT 427
QY 211 CGGCGCGCTCGCGGAGTGGCTTCCGCGCTCGGCGCTCGG 244
DB 428 CGGCGTGGCGCGCGCGCAGCGCGCGCGGCTCGG 461

RESULT 12
LOCUS CC626818 839 bp DNA linear GSS 19-JUN-2003
DEFINITION OGPARE2TH ZM 0.7.1.5 KB Zea mays genomic clone ZMEMA0354K04,
ACCESSION CC626818
VERSION CC626818
KEYWORDS genomic survey sequence.
SOURCE GSS.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 839)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGPARE2TV
Contact: Cathy Whitelaw

ORGANISM TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seg primer: TR
Class: sheared ends.
Location/Qualifiers
1. 839
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMEMA0354K04"
/clone_1ib="ZM_0.7.1.5_KB"
/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN
Query Match 16.7%; Score 41.2; DB 9; Length 839;
Best Local Similarity 49.5%; Pred. No. 4.4;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 31 CAATCCGAGCATGACGAGCAACAGCAGGTGAGTGGCTGTGACCCCTGACACAG 90
DB 581 CGACGCGCGGCTACGCGCGCGCGCGCTCCGACGCGCGGTAGCGCCCGACAG 640
QY 91 GCGCCGATGAGGTTCAATTCCTCATGTCACCCGTACAGAGCCCTCCACTCGGA 150
DB 641 CGTCCAGGCGGAGCAGCAGCGGCATGAGCGCTCGGAGAACCGGCGCGCTCGT 700
QY 151 GGGGGCTTGGCGCTTCTGAGGGTTCCGCGTCAAGCGCGCTCGCGCTGGGGAGCT 210
DB 701 CGCCGCGGACCGCTCCACGCTCTCTGCTCCGCGGAGAGCGCTCCAGCGCGCTGCT 760
QY 211 CGGCGCGCTCGCGGAGTGGCTTCCGCGCTCGGCGCTCGG 244
DB 761 CGGCGTGGCGCGCGCGCAGCGCGCGGCGGTGCG 794

RESULT 13
LOCUS CNS0091P 925 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence RT3 end of BAC # BACR19D16 of RPL1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL053013
VERSION AL053013
KEYWORDS
SOURCE GSS.
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPL1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or

```
library(MPMGP637")
/ote="Vector: pSport1; Site 1: NotI; Site 2: SalI;
oligo-dt=NotI primed, SalI adaptors, directionally c
library normalised by oligonucleotide fingerprinting
```

16.6%; Score 41; DB 4; Length 738;

Sequence	Identity	57.4%	Pred. No. 5;	Mismatches	55;	Indels	0;	Gaps
GTATACGAAAGCCCCCTCCACTCGAGAGGGGCTTCGGCGTTCGAGGGTTC								

1

QY 189 GTGGGCTGGGCGCTGGGGGACTCGGCCCGTGGGCGGAGTGGCTTGGC 238
Db 141 GGCCTGCACAGGCGGGCGGAGAGGCTGAGAGCGGCGGCGCTTGGCC 190

Search completed: April 15, 2005, 17:41:36
Job time : 3303 secs

Mon Apr 18 09:47:33 2005

us-09-855-340b-6.rmpb

Pa

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 17:44:42 ; Search time 514 Seconds
(without alignments)
2915.326 Million cell updates/sec

Title: US-09-855-340B-6

Perfect score: 247
Sequence: 1 taggggaatccacccgcggag.....gtgcctccggtccgggga 247

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	247	100.0	247 9 US-09-855-340-6	Sequence 6, Appli
2	147	59.5	241 9 US-09-855-340-4	Sequence 4, Appli
3	124	50.2	260 9 US-09-855-340-9	Sequence 9, Appli
4	44.2	17.9	624 18 US-10-425-115-73105	Sequence 17794, A
5	42	17.0	1232 18 US-10-437-963-17794	Sequence 17794, A
6	41.4	16.8	394 18 US-10-425-115-31984	Sequence 31984, A
7	41.2	16.7	801 18 US-10-425-115-35614	Sequence 35614, A
8	40	16.2	1446 17 US-10-369-493-42829	Sequence 42829, A
9	40	16.2	1767 15 US-10-156-761-4150	Sequence 4150, Ap
10	40	16.2	9025608 15 US-10-156-761-1	Sequence 1, Appli
11	39.8	16.1	813 15 US-10-156-761-5740	Sequence 5740, Ap

12	39.2	15.9	1212 15 US-10-156-761-3345	Sequence 334
13	39.2	15.9	9025608 15 US-10-156-761-1	Sequence 1,
14	38.6	15.6	512 18 US-10-437-963-30223	Sequence 3022
15	37.8	15.3	2082 15 US-10-156-761-906	Sequence 906
16	37.6	15.2	654 18 US-10-767-701-25152	Sequence 2515
17	37.6	15.2	662 18 US-10-363-345A-639	Sequence 639A
18	37.6	15.2	662 18 US-10-363-345A-640	Sequence 640A
19	37.6	15.2	662 19 US-10-363-483A-639	Sequence 639A
20	37.6	15.2	662 19 US-10-363-483A-640	Sequence 640A
21	37.4	15.1	1032 15 US-10-156-761-4218	Sequence 4218
22	37.4	15.1	5786 17 US-10-188-832-11	Sequence 11,
23	37.4	15.1	418550 17 US-10-292-798-1463	Sequence 1463
24	37.2	15.1	2574 18 US-10-437-963-18540	Sequence 1854
25	37.2	15.0	942 18 US-10-437-963-15838	Sequence 1583
26	37	15.0	1239 17 US-10-260-238-1271	Sequence 1271
27	37	15.0	1284 18 US-10-437-963-53046	Sequence 5304
28	37	15.0	1637 17 US-10-260-238-1233	Sequence 1233
29	37	15.0	3834 15 US-10-156-761-2681	Sequence 2681
30	37	15.0	4065 10 US-09-791-254-1	Sequence 1,
31	36.8	14.9	402 18 US-10-437-963-46295	Sequence 46
32	36.8	14.9	758 18 US-10-767-701-11567	Sequence 11,
33	36.6	14.8	1242 17 US-10-302-172-266	Sequence 26
34	36.6	14.8	1452 15 US-10-156-761-6328	Sequence 63
35	36.6	14.8	7536 10 US-09-940-316B-1	Sequence 1,
36	36.2	14.7	588 18 US-10-437-963-39473	Sequence 39
37	36.2	14.7	937 17 US-10-437-963-97455	Sequence 97
38	36	14.6	927 17 US-10-282-122A-26176	Sequence 26
39	36	14.6	1161 17 US-10-282-122A-28713	Sequence 28
40	36	14.6	1436 18 US-10-767-701-13683	Sequence 13
41	36	14.6	1446 15 US-10-156-761-4324	Sequence 43
42	35.8	14.5	861 15 US-10-437-963-94786	Sequence 94
43	35.8	14.5	831 15 US-10-156-761-5578	Sequence 55
44	35.8	14.5	1242 18 US-10-437-963-28275	Sequence 28
45	35.6	14.4	627 18 US-10-437-963-42823	Sequence 42

ALIGNMENTS

RESULT 1
US-09-855-340-6
Sequence 6, Application US/09855340
Patent No. US20020076788A1
GENERAL INFORMATION:
APPLICANT: Hosted, Jr., Thomas J.
TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africa
TITLE OF INVENTION: PMPI integrase and use of integrating function
TITLE OF INVENTION: site-specific integration into Micromonospora
TITLE OF INVENTION: halophilica and Micromonospora chromoc
FILE REFERENCE: INO1164K
CURRENT APPLICATION NUMBER: US/09/855,340
PRIORITY FILING DATE: 2001-05-15
PRIORITY FILING DATE: 2000-05-17
PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 247
TYPE: DNA
ORGANISM: Micromonospora carbonacea

Query Match 100.0%; Score 247; DB 9; length 247;
Best Local Similarity 100.0%; Pred. No. 2.5e-63;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 TAGGGGAATCCACTCCGAGAGACGCCCGAGACATCCGAGAGTGAACGAGCAACGAG
DB 1 TAGGGGAATCCACTCCGAGAGACGCCCGAGACATCCGAGAGTGAACGAGCAACGAG
QY 61 GTGAGGTGCTCTTGACACCCCTGACGAGGCGCGGTACGAGGTTCATTCATCA

Query Match	16.8%	Score 41.4;	DB 18;	Length 394;
Best Local Similarity	48.1%;	Pred. No. 0.011;		
Matches 102;	Conservative	0;	Mismatches 110;	Indels 0;
				Gaps 0;

```

RESULT 8
US-10-369-493-42829/C
; Sequence 42829, Application US/10369493
; Publication No. US20030233675A1
GENERAL INFORMATION:
;
; APPLICANT: Cao, Yongwei
;
; APPLICANT: Hinkle, Gregory J.
;
; APPLICANT: Slater, Steven C.
;
; APPLICANT: Goldman, Barry S.
;
; APPLICANT: Chen, Xianfeng
;
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR P

```

;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 42829
;; LENGTH: 1446
;; TYPE: DNA
;; ORGANISM: Myxococcus xanthus
US-10-369-493-42829

Query Match 16.2%; Score 40; DB 17; Length 1446;
Best Local Similarity 59.8%; Pred. No. 0.024;
Matches 67; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 133 AGCCCCCTCCACCTCGAGAGGGGCTTCGCGGCTTCGAGGGTTCGCGGCTCGAGGGCTCG 192
DB 1227 AGCGCTTCCTCGGCGGAGCTGCGGCTTCGCGGCTTCGCGGAGCCCAACGAGCTCGGCGCGG 1168
QY 193 GCTCGGCGCTGGGGAGCTCGGCGGCTCGGCGGAGTGGGCTCGGCTCGCGG 244
DB 1167 GAGCGCGGCGGCTCGACCTTCGCGGCTCGGCGGAGCGCAAGGCTTCAGG 1116

RESULT 9

US-10-156-761-4150/c
;; Sequence 4150, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OMURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIBA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 4150
;; LENGTH: 1767
;; TYPE: DNA
;; ORGANISM: Streptomyces avermitilis
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)..(1767)
US-10-156-761-4150

Query Match 16.2%; Score 40; DB 15; Length 1767;
Best Local Similarity 53.1%; Pred. No. 0.024;
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 84 GACCGAGGCGCCCGGTAACGGTTCATTCACATGATCACCCGTAACGAGGCGCCCTCC 143
DB 1253 GACCCGAGCTGGAGCGGCTTCGCTGCGGTGCGAGCGGCTTCGCTGCGGCTCCG 1194
QY 144 ACTCGAGAGGGGAGCTTCGCGGCTTCGAGGGTTCGCGGCTTCGAGGGTTCGCGGCTG 203
DB 1193 GTCTCGGCGGCGGCTTCGAGAGCGGCTTCGAGGCGGCTTCGAGGCTTCGAGGCGG 1134
QY 204 GGGGACTCGGCGCCGCTCGCGGCGGAGTGGGCTTCGCGGCTCG 243
DB 1133 GTGCGGCGGCTTCGAGGCTTCGCGGAGCTTCGCGGCTCG 1094

RESULT 10
US-10-156-761-1
;; Sequence 1, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OMURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIBA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 1
;; LENGTH: 9025608
;; TYPE: DNA
;; ORGANISM: Streptomyces avermitilis
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (4187715)
;; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 16.2%; Score 40; DB 15; Length 9025608;
Best Local Similarity 53.1%; Pred. No. 0.011;
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 84 GACCGAGGCGCCCGGTAACGGTTCATTCACATGATCACCCGTAACGAGGCGCCCTCC 143
DB 5115914 GCGCCGAGCTGGAGCGGCTTCGCTGCGGTGCGAGCGGCTTCGCGGCTCCG 5115973
QY 144 ACTCGAGAGGGGAGCTTCGCGGCTTCGAGGGTTCGCGGCTTCGAGGGTTCGCGGCTG 203
DB 5115974 GTCTCGGCGGCGGCTTCGAGAGCGGCTTCGAGGCGGCTTCGAGGCGGCTTCGAGGCGG 5116033
QY 204 GGGGACTCGGCGCCGCTCGCGGCGGAGTGGGCTTCGCGGCTCG 243
DB 5116034 GTGCGGCGGCTTCGAGGCTTCGCGGAGACTTCGCGGCTCG 5116073

RESULT 11

US-10-156-761-5740/c
;; Sequence 5740, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OMURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIBA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 5740
;; LENGTH: 813
;; TYPE: DNA
;; ORGANISM: Streptomyces avermitilis

Mon Apr 18 09:47:33 2005

us-09-855-340b-6.rmp

FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(813)
US-10-156-761-5740

Query Match
Best Local Similarity 57.7%; Score 39.8; DB 15; Length 813;
Pred. No. 0.029;
Matches 71; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 124 CGTACGAGAGGCCCCCTTCATCGAGGGGCGCTTCGCGCTTCCTGAGGGTTCCGGTC 183
DB 265 CGCGACGAGAGGCGCGCTTCGAGTAGTAGGGGCTTCCTGCTTCCTTCGAGGGTCC 206
QY 184 AGCGGCTGCGCTTCGCGCTTCGAGGGGACTCGGCCCCCTTCGCGGGAGTGGCTCGCGTCCG 243
DB 205 AGACGCGCGCGCGCGCGGGGGCGCGCTTCCTTCGCTGCGCTGTGGGGGCTTCAGTGG 146
QY 244 GGG 246
DB 145 CGG 143

RESULT 12
US-10-156-761-3345
Sequence 3345, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 3345
LENGTH: 1212
TYPE: DNA
ORGANISM: Streptomyces avermitilis

FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1212)
US-10-156-761-3345

Query Match
Best Local Similarity 53.2%; Score 39.2; DB 15; Length 1212;
Pred. No. 0.043;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 90 GGGCCCCGTACGGGTTCAATCCCATCACTACACCCGTACAGAGGCCCCCTTCCTACTCGG 149
DB 641 GGCACCTCGACGGCTTCGATCACTCAAGCGAGAAACAACACTTCAGATCCACGCGG 700
QY 150 AGGGGCTTCGCGCTTCCTGAGGGTTTCGCGGTCAAGCGGCTCGGCTCGGCGCTGGGGGAC 209
DB 701 GGGAGGCTTCGCTTCCTGATCTGGCAGGCGCTTCAGTGTGTGGGCGGACCGG 760
QY 210 TCGGCCCGCTCGCGGAGTGGCTTCGCGGCTCGGCGG 245
DB 761 TCGGTACGCGCGTGGCATCATGACGACATCCAGG 796

RESULT 13
US-10-156-761-1/c
Sequence 1, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1

LENGTH: 9025608

TYPE: DNA

ORGANISM: Streptomyces avermitilis

FEATURE:

NAME/KEY: misc feature

LOCATION: (4187715)

OTHER INFORMATION: a, t, c, g, other or unknown

US-10-156-761-1

Query Match
Best Local Similarity 53.2%; Score 39.2; DB 15; Length 9025608;
Pred. No. 0.019;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 90 GGGCCCCGTACGGGTTCAATCCCATCACTACACCCGTACAGAGGCCCCCTTCCTACTCGG 149
DB 4173750 GGCACCTCGACGGCTTCGATCACTCAAGCGGAGAAACAACACTTCAGATCCACG 700
QY 150 AGGGGCTTCGCGCTTCCTGAGGGTTTCGCGGTCAAGCGGCTCGGCTCGGCGCTGGGGGAC 209
DB 4173690 GGGAGGCTTCGCTTCCTGATCTGGCAGGCGCTTCAGTGTGTGGGCGGACCG 760
QY 210 TCGGCCCGCTCGCGGAGTGGCTTCGCGGCTCGGCGG 245
DB 4173630 TCGGTACGCGCGTGGCATCATGACGACATCCAGG 796

RESULT 14
US-10-437-963-30223/c
Sequence 30223, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: LA ROSA, THOMAS J.
APPLICANT: KOVALLIC, DAVID K.
APPLICANT: ZHOU, YIHUA
APPLICANT: CAO, YONGWEI
APPLICANT: WU, WEI
APPLICANT: BOUKHAROV, ANDREY A.
APPLICANT: BARBAZUK, BRAD
APPLICANT: LI, PING
TITLE OF INVENTION: RICE NUCLEIC ACID MOLECULES AND OTHER MOLECULES
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 30223
LENGTH: 512
TYPE: DNA
ORGANISM: Oryza sativa

FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(512)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_34649C.1

US-10-437-963-30223

Query Match 15.6%; Score 38.6; DB 18; Length 512;
Best Local Similarity 46.4%; Pred. No. 0.069;
Matches 83; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 68 GGCCTGTTGACCCCTGACGAGGCCCCGAGGAGTTCAATCCATGAGTACACCCGTA 127
DB 472 GNCGCGTGNCCNCCGCGGCGCNCCTGCGGCTCGCCGCTTTTNCGTA 413
QY 128 CACGAGAGGCCCCCTCACTGAGGAGGCGCTTTCGAGGTTTCGCGTCAAGC 187
DB 412 CTCCTGCCCCCGCCGCGGCTCCGCTGAGGTCNTGCTGAGGAGGAGG 353
QY 188 GGTCCGCTGCGGCTGAGGAGTCCGCCCCGTCGCGGAGTGCCTCGCGTCCGCGG 246
DB 352 GGTGAGNCCGCGCGTGCAGCCGTCGCGGCGCGGATGAGTGTGTCGAGG 294

RESULT 15

US-10-156-761-906/C
; Sequence 906, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 906
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2082)
US-10-156-761-906

Query Match 15.3%; Score 37.8; DB 15; Length 2082;
Best Local Similarity 57.0%; Pred. No. 0.11;
Matches 69; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 123 CCGTACACGAGCCCTCTCACTCGAGGAGGCGCTTCGCGCTTCTGAGGGTTCCGGT 182
DB 467 CCCTGACGACGTCACCTGTCGCGGAGATGATGTGCTCCGCTCGCACACGACAGTCA 408
QY 183 CAGGCGGTGCTGCGGCTGAGGAGTCCGCGGAGTGCCTTCGCGGTCC 242
DB 407 CCGACGCGAGCTCGGTGCGGAGACACGATCCGTCGCGCCCTGCGGCGCACGAGTTC 348
QY 243 G 243
DB 347 G 347

Search completed: April 15, 2005, 19:10:14
Job time : 539 secs


```

: sequence 2, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: . LENGTH: 4403765
: . TYPE: DNA
: . ORGANISM: Mycobacterium tuberculosis
: . FEATURE:
: . OTHER INFORMATION: CDC 1551
: . OTHER INFORMATION: "n" bases at various positions throughout the sequence
: US-09-103-840A-2

```


Query Match 14.9%; Score 36.8; DB 3; Length 4403765;
 Best Local Similarity 48.6%; Pred. No. 4;
 Matches 101; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

39 AGCATGACGAGCAACCAAGTCAAGTGGCTTTGACCCCTTGACCAAGGCGCCGCT 98
 Db 3631715 AGCACCAGGAGCGGACCGCCCTGCTGTCCCGAGCGACCTCATACGCTGCCACACG 3631774
 99 ACGGTTCAATTCATCATGTAACCCGTACAGAGGCCCTTCCACTGAGAGGGGGCT 158
 Db 3631775 ACAACCCGCGCCCTTGGGCAACCCGATTCGGCGCGCCCGCCGACCAAGCGCGCTGCG 3631834
 159 TCGGCGTTCTTGAAGGTTGCGGTCAAGCGGTGCGTCCGCGCTGGGGGACTCGCCCG 218
 Db 3631835 GGGTGTGTCGCGCAACATCAAGCAAGCGCGCCGACCAAGCGCGCGCGCTG 3631894
 219 TCGGCGGAGTGCGCTTCGCGCTCGGGG 246
 Db 3631895 GCGAGCAGATCGGCTCGGCGCGCCCG 3631922

RESULT 6

US-09-103-840A-1

Sequence 1, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

TITLE OF INVENTION: TUBERCULOSIS

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103, 840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 1

LENGTH: 4411529

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match 14.9%; Score 36.8; DB 3; Length 4411529;
 Best Local Similarity 48.6%; Pred. No. 4;
 Matches 101; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

39 AGCATGACGAGCAACCAAGTCAAGTGGCTTTGACCCCTTGACCAAGGCGCCGCT 98
 Db 3636866 AGCACCAGGAGCGGACCGCCCTGCTGTCCCGAGCGACCTCATACGCTGCCACACG 3636925
 99 ACGGTTCAATTCATCATGTAACCCGTACAGAGGCCCTTCCACTGAGAGGGGGCT 158
 Db 3636926 ACAACCCGCGCCCTTGGGCAACCCGATTCGGCGCGCCCGCCGACCAAGCGCGCTGCG 3636985
 159 TCGGCGTTCTTGAAGGTTGCGGTCAAGCGGTGCGTCCGCGCTGGGGGACTCGGCCCG 218
 Db 3636986 GGGTGTGTCGCGCAACATCAAGCAAGCGCGCCGACCAAGCGCGCGCTG 3637045
 219 TCGGCGGAGTGCGCTTCGCGCTCGGGG 246
 Db 3637046 GCGAGCAGATCGGCTCGGCGCGCCCG 3637073

RESULT 7

US-09-799-451-266/C

Sequence 266, Application US/09799451

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Kyle

APPLICANT: Asundi, Vinod

APPLICANT: Ren, Feiyang

APPLICANT: Zhang, Jie

APPLICANT: Xue, Aidong J.

APPLICANT: Zhao, Qing A.

APPLICANT: Wang, Jian-Rui

APPLICANT: Ma, Yundong

APPLICANT: Yamazaki, Victoria

APPLICANT: Chen, Rui-hong

APPLICANT: Wang, Zhiwei

APPLICANT: Wang, Dunrui

APPLICANT: Yang, Yonghong

APPLICANT: Wehrman, Tom

APPLICANT: Ghosh, Reena

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 803

CURRENT APPLICATION NUMBER: US/09/799,451

CURRENT FILING DATE: 2001-03-05

NUMBER OF SEQ ID NOS: 948

SOFTWARE: PC_FU_genes Version 2.0

SEQ ID NO 266

LENGTH: 1342

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(786)

NAME/KEY: misc feature

LOCATION: (1)..(1342)

OTHER INFORMATION: n = a, t, c or g

US-09-799-451-266

Query Match 14.8%; Score 36.6; DB 4; Length 1342;
 Best Local Similarity 47.6%; Pred. No. 1.2;
 Matches 108; Conservative 0; Mismatches 119; Indels 0; Gap

20 GAGCCCGGAGCAATCCGAGCATGACGAGCAACCAAGTCAAGTGGCTTTG 111
 Db 287 GCGAGGAGGAGGAGTGGCTGAGGCGCGCCGCGCAACGAGCACTTCCAGTGA 111
 80 CCTGACCAAGGCGCCCGGTACAGGTTCAATTCATCATGTAACCCGTACAGAGGCC 111
 Db 227 CGTGGGAGGAGCTCCAGGCGCCCGCTACAGCGCCCGCACTTCTTACAGCAGAGCTCC 111
 140 CTCACCTCGAGAGGCGGCTTTCGAGGTTCTGAGGTTTCGCGTCAAGCGCTCGCTCC 111
 Db 167 GCGTGAAGCTTGAAGGCTGCGCGGCTTTCGAGGTTTCGCGTCAAGCGCTCGCTCC 111
 200 GCTGGGAGGAGTGGCGCCCGGTGAGGAGTGGCTTTCGAGGTTTCGCGTCAAGCGCTCG 111
 Db 107 GATGGCAGCAGCGGCGGCGGAGGAGCAGTGTGCGCGCGCGCGCGCGCG 61

RESULT 8

US-09-410-551B-1

Sequence 1, Application US/09410551B

Patent No. 6503737

GENERAL INFORMATION:

APPLICANT: KOSAN BIOSCIENCES, Inc.

APPLICANT: REEVES, CHRISTOPHER

APPLICANT: CHU, DANIEL

APPLICANT: KHOSLA, CHAITAN

APPLICANT: SANTI, DANIEL

APPLICANT: WU, KAI

TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA

TITLE OF INVENTION: CONSTRUCTS THEREFOR

FILE REFERENCE: 30062-20026.00

CURRENT APPLICATION NUMBER: US/09/410, 551B

CURRENT FILING DATE: 1999-10-01

PRIOR APPLICATION NUMBER: US 60/139, 650

PRIOR FILING DATE: 1999-06-17

Mon Apr 18 09:47:33 2005

us-09-855-340b-6.rn1

Db 330 CGCGGA 325

RESULT 11
US-09-283-471A-38/c
; Sequence 38, Application US/09283471A
; Patent No. 6340673
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method For Treating Tumorigenic Diseases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/283,471A
; FILING DATE: 04-APR-1999
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,533
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 27373/32742A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-283-471A-38

Query Match 14.7%; Score 36.4; DB 3; Length 1280;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 70; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Db 122 CCCGTACAGAGCCCTTCACCTCGAGAGGGGCTTCGCGCTTCCTGAGGGTTCCGCG 181
|||
450 CCCCCCCCCCG 391
|||
QY 182 TCAGGCGGTGCGTCCG 241
|||
Db 390 CTCGTGCG 331
|||
QY 242 CGGCGGA 247
|||
Db 330 CGGCGGA 325

US-08-483-533-39/c
; Sequence 39, Application US/08483533
; Patent No. 6172047
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 28097/32742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-483-533-39

Query Match 14.7%; Score 36.4; DB 3; Length 1300;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 70; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Db 122 CCCGTACAGAGCCCTTCACCTCGAGAGGGGCTTCGCGCTTCCTGAGGGTTCCGCG 470
|||
QY 182 TCAGGCGGTGCGTCCG 410
|||
Db 410 CTCGTGCG 242
|||
QY 242 CGGCGGA 247
|||
Db 350 CGGCGGA 345

RESULT 13
US-09-283-471A-39/c
; Sequence 39, Application US/09283471A
; Patent No. 6340673
; GENERAL INFORMATION:

NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/283,471A
FILING DATE: 04-APR-1999
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,533
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 27373/32742A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1300 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-283-471A-39
Query Match 14.7%; Score 36.4; DB 3; Length 1300;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 70; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 122 CCGGTACAGAAAGCCCTTCACTCGAGAGGGGCGTTCCTGAGAGGTTCCGG 181
DB 470 CCGGCCCCCTGGGCGCGCGCGGTGGGGGCGGCGCGGTGGGCGCGGCGG 411
QY 182 TCAGCGCGTGGCTCGGCGCTGGGGAAGTCCGCCCGTGGCGGAGTGGCTCGGCGTC 241
DB 410 CTCGTGGCGCGCGCTCGGCGGCGGCGGTCTCGGCGCAGTCTGTCATGTCGTGCGGA 351
QY 242 CGGCGA 247
DB 350 CGCGGA 345
RESULT 14
US-08-483-533-36/c
Sequence 36, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
DISEASES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-483-533-36
Query Match 14.7%; Score 36.4; DB 3; Length 1327;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 70; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 122 CCGGTACAGAAAGCCCTTCACTCGAGAGGGGCGTTCCTGAGAGGTTCCGG 181
DB 470 CCGGCCCCCTGGGCGCGCGCGGTGGGGGCGGCGCGGTGGGCGCGGCGG 411
QY 182 TCAGCGCGTGGCTCGGCGCTGGGGAAGTCCGCCCGTGGCGGAGTGGCTCGGCGTC 241
DB 410 CTCGTGGCGCGCGCTCGGCGGCGGCGGTCTCGGCGCAGTCTGTCATGTCGTGCGGA 351
QY 242 CGGCGA 247
DB 350 CGCGGA 345

RESULT 15
US-09-283-471A-36/c
Sequence 36, Application US/09283471A
Patent No. 6340673
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method For Treating Tumorigenic Diseases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 19:01:23 ; Search time 871.841 Seconds
(without alignments)
8204.076 Million cell updates/sec

Title: US-09-855-340B-1
Perfect score: 1 gctgcgacgcgaagaacg.....tctgcgcgcagcagcatga 1179
Sequence: 1 gctgcgacgcgaagaacg.....tctgcgcgcagcagcatga 1179

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 5622541 seqs, 303355566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09E_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1179	100.0	1179	9	US-09-855-340-1
2	1176	99.7	1176	11	US-09-758-759-176
3	65.8	5.6	1137	15	US-10-156-761-3714
4	65.8	5.6	9025608	15	US-10-156-761-1
5	57.2	4.9	1743	17	US-10-369-493-4111
6	56.6	4.8	1161	9	US-09-821-167-15
7	56.6	4.8	11188	9	US-10-156-761-3695
8	55.8	4.7	1179	15	US-10-203-295-2
9	55.8	4.7	27541	18	US-10-203-295-35
10	55.8	4.7	125401	18	US-10-203-295-35
11	54.4	4.6	846	17	US-10-260-238-5339

12	53.6	4.5	879	15	US-10-156-761-5080
13	53.2	4.5	1104	19	US-10-762-107-45
14	53.2	4.5	36602	19	US-10-762-107-1
15	52.2	4.4	426	15	US-10-156-761-5794
16	52	4.4	86421	9	US-09-976-059-1
17	51.6	4.4	3363	19	US-10-211-028-135
18	51.6	4.4	23949	19	US-10-211-028-106
19	51.4	4.4	516	9	US-09-874-923-34
20	51.4	4.4	582	9	US-09-991-496-34
21	51.4	4.4	582	9	US-09-991-496-114
22	51.4	4.4	7065	9	US-09-991-496-115
23	51.4	4.4	7065	9	US-09-991-496-115
24	50.8	4.3	1167	15	US-10-156-761-2743
25	50.8	4.3	2136	17	US-10-282-122A-25513
26	50.8	4.3	2771	18	US-10-425-115-81904
27	50.6	4.3	1659	15	US-10-156-761-6116
28	50.4	4.3	1539	15	US-10-156-761-75
29	50.4	4.3	11238	15	US-10-205-032-15
30	50.4	4.3	60196	15	US-10-205-032-1
31	50.4	4.3	9025608	15	US-10-156-761-1
32	50.2	4.3	1914	18	US-10-425-115-142240
33	50.2	4.2	1647	17	US-10-369-493-28531
34	49.8	4.2	1647	17	US-10-369-493-31290
35	49.8	4.2	1923	15	US-10-156-761-1823
36	49.8	4.2	2073	11	US-09-758-759-158
37	49.8	4.2	109519	11	US-09-758-759-1
38	49.6	4.2	135638	16	US-10-314-963-42144
39	49.6	4.2	1476	18	US-10-156-761-7069
40	49.4	4.2	714	15	US-10-156-761-860-846
41	49.2	4.2	3244	15	US-10-723-860-846
42	49	4.2	3244	15	US-10-084-817-320
43	48.8	4.1	984	18	US-10-767-701-10839
44	48.8	4.1	1437	15	US-10-156-761-5388
45	48.8	4.1	1437	15	US-10-156-761-5388

ALIGNMENTS

RESULT 1
US-09-855-340-1
Sequence 1, Application US/09855340
Patent No. US20020076788A1
GENERAL INFORMATION:
APPLICANT: Hosted, Jr., Thomas J.
TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africa
TITLE OF INVENTION: MPEI integrase and use of integrating function
TITLE OF INVENTION: site-specific integration into Micromonospora
TITLE OF INVENTION: halophilica and Micromonospora chromoc
FILE REFERENCE: IN01164K
CURRENT APPLICATION NUMBER: US/09/855,340
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/204,670
PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1179
TYPE: DNA
ORGANISM: Micromonospora carbonacea
US-09-855-340-1
Query Match
Best Local Similarity 100.0%; Score 1179; DB 9;
Matches 1179; Conservative 0; Mismatches 0; Indels 0; Gap

1 GTGTGATCGAGAGAGAGCGGCGCTTACCGCATTCGGGACCTCGGTGCGGTAA
1 GTGTGATCGAGAGAGAGCGGCGCTTACCGCATTCGGGACCTCGGTGCGGTAA
61 GTGTGATCGAGAGAGAGCGGCGCTTACCGCATTCGGGACCTCGGTGCGGTGCGG

Db 61 GTGACCATTCAGACCGGTTATCCGACGAGACCGGCGCAAGATGCGATGTCAGTTTC 120
QY 121 CGTGGGAGAGATTGAGGGGCAACGGCTATGCGCGGCGGCGTACGATTACCTCGCC 180
Db 121 CGTGGGAGAGATTGAGGGGCAACGGCTATGCGCGGCGGCGTACGATTACCTCGCC 180
QY 181 GATTGTTGGGGAGTGTGGCGAGCTACGAAAGACGCTGAAACCGACCGCGTGAAC 240
Db 181 GATTGTTGGGGAGTGTGGCGAGCTACGAAAGACGCTGAAACCGACCGCGTGAAC 240
QY 241 TCGGAGGAGCAACCGGATCCGACCAACCTCTGCTCCATATCTCGGCGCATCTCACTTAC 300
Db 241 TCGGAGGAGCAACCGGATCCGACCAACCTCTGCTCCATATCTCGGCGCATCTTACCTTGAAC 300
QY 301 GAGCTGGAAGGGGAGGTACCCAGACGATGGGTCAAGAGCTTGAAGGCGGCGTCCGCGCG 360
Db 301 GAGCTGGAAGGGGAGGTACCCAGACGATGGGTCAAGAGCTTGAAGGCGGCGTCCGCGCG 360
QY 361 TGGCCGAGATCCACCGGGGGTGTGGGAGAGCGGCTGGCGGCAACGATCAGCAATGTC 420
Db 361 TGGCCGAGATCCACCGGGGGTGTGGGAGAGCGGCTGGCGGCAACGATCAGCAATGTC 420
QY 421 CACGGCTCTGTCACACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
Db 421 CACGGCTCTGTCACACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
QY 481 CCGTCTCTTTCAGACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
Db 481 CCGTCTCTTTCAGACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
QY 541 CCGGAGATCGGCTCGGCTTATACGCGCGCTTCGCGGCACTGCGGCAACGATCAGCTCAAC 600
Db 541 CCGGAGATCGGCTCGGCTTATACGCGCGCTTCGCGGCACTGCGGCAACGATCAGCTCAAC 600
QY 601 CTGCTGCGGAGACCGGCTCTGAGGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 660
Db 601 CTGCTGCGGAGACCGGCTCTGAGGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 660
QY 661 CTGCTGCGGAGACCGGCTCTGAGGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 720
Db 661 CTGCTGCGGAGACCGGCTCTGAGGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 720
QY 721 GGAAGAGCTCTTTCAGTCTGAGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 780
Db 721 GGAAGAGCTCTTTCAGTCTGAGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 780
QY 781 AGGAAAGTCTCTTTCAGTCTGAGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 840
Db 781 AGGAAAGTCTCTTTCAGTCTGAGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 840
QY 841 TTTACCGGCGGCAACCGGAGTGTGAGGCGGCAATTTCCGCGGAGTCTGGGCTGAC 900
Db 841 TTTACCGGCGGCAACCGGAGTGTGAGGCGGCAATTTCCGCGGAGTCTGGGCTGAC 900
QY 901 GCGTGGAGGAGGCGGAGTCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 960
Db 901 GCGTGGAGGAGGCGGAGTCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 960
QY 961 GCGATCTGATTTTCGCGGCGGCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 1020
Db 961 GCGATCTGATTTTCGCGGCGGCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 1020
QY 1021 TCGATTCGCGGAGTCTGCTGTAAGGCGGCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 1080
Db 1021 TCGATTCGCGGAGTCTGCTGTAAGGCGGCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 1080
QY 1081 CTCGCGGAGATGAGAGGCGGCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 1140
Db 1081 CTCGCGGAGATGAGAGGCGGCTGCGGCTTACGATTCAGATTTGCGGCGGCACTCACTGAC 1140
QY 1141 GACGAGAGCTGACGAGCTGTGTCGCGGCAACGATGAT 1179
Db 1141 GACGAGAGCTGACGAGCTGTGTCGCGGCAACGATGAT 1179
QY 1141 GACGAGAGCTGACGAGCTGTGTCGCGGCAACGATGAT 1179
Db 1141 GACGAGAGCTGACGAGCTGTGTCGCGGCAACGATGAT 1179

RESULT 2
US-09-758-759-176
Sequence 176, Application US/09758759
Publication No. US20040101832A1
GENERAL INFORMATION:
APPLICANT: Hosted, Thomas J.
APPLICANT: Wang, Tim X.
APPLICANT: Horan, Ann C.
FILE OF INVENTION: Evernimolcin Biosynthetic Genes
FILE REFERENCE: ID0983K US
CURRENT APPLICATION NUMBER: US/09/758,759
CURRENT FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: US 60/175,751
PRIOR FILING DATE: 2000-01-12
NUMBER OF SEQ ID NOS: 204
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 176
LENGTH: 1176
TYPE: DNA
ORGANISM: Micromonospora carbonacea
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1176)
US-09-758-759-176

Query Match 99.7%; Score 1176; DB 11; Length 1176;
Best Local Similarity 100.0%; Pred. No. 1e-301;
Matches 1176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTGATTCGAGAGAGAGGAGGCGGCTTACCGGATTCGGGAGCTCTGCGGCTGAAAG 60
Db 1 GTGTGATTCGAGAGAGAGGAGGCGGCTTACCGGATTCGGGAGCTCTGCGGCTGAAAG 60
QY 61 GTGACCATTCAGACCGGTTATCCGACGAGACCGGCGCAAGATGCGATGTCAGTTTC 120
Db 61 GTGACCATTCAGACCGGTTATCCGACGAGACCGGCGCAAGATGCGATGTCAGTTTC 120
QY 121 CGTGGGAGAGATTGAGGGGCAACGGCTATGCGCGGCGGCGTACGATTACCTCGCC 180
Db 121 CGTGGGAGAGATTGAGGGGCAACGGCTATGCGCGGCGGCGTACGATTACCTCGCC 180
QY 181 GATTGTTGGGGAGTGTGGCGAGCTACGAAAGACGCTGAAACCGACCGCGTGAAC 240
Db 181 GATTGTTGGGGAGTGTGGCGAGCTACGAAAGACGCTGAAACCGACCGCGTGAAC 240
QY 241 TCGGAGGAGCAACCGGATCCGACCAACCTCTGCTCCATATCTCGGCGCATCTCACTTGAAC 300
Db 241 TCGGAGGAGCAACCGGATCCGACCAACCTCTGCTCCATATCTCGGCGCATCTCACTTGAAC 300
QY 301 GAGCTGGAAGGGGAGGTACCCAGACGATGGGTCAAGAGCTTGAAGGCGGCGTCCGCGCG 360
Db 301 GAGCTGGAAGGGGAGGTACCCAGACGATGGGTCAAGAGCTTGAAGGCGGCGTCCGCGCG 360
QY 361 TGGCCGAGATCCACCGGGGGTGTGGGAGAGCGGCTGGCGGCAACGATCAGCAATGTC 420
Db 361 TGGCCGAGATCCACCGGGGGTGTGGGAGAGCGGCTGGCGGCAACGATCAGCAATGTC 420
QY 421 CACGGCTCTGTCACACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
Db 421 CACGGCTCTGTCACACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
QY 481 CCGTCTCTTTCAGACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
Db 481 CCGTCTCTTTCAGACGATCTGCGGCGGCGGATCGGCGGCAACGATCAGCTCAAC 480
QY 541 CCGGAGATCGGCTCGGCTTATACGCGCGCTTCGCGGCACTGCGGCAACGATCAGCTCAAC 600
Db 541 CCGGAGATCGGCTCGGCTTATACGCGCGCTTCGCGGCACTGCGGCAACGATCAGCTCAAC 600
QY 601 CTGCTGCGGAGACCGGCTCTGAGGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 660
Db 601 CTGCTGCGGAGACCGGCTCTGAGGTGGGTGAGGCGATCGGCTGCGGCGGCGGCGGCTGAC 660


```

RESULT 3
US-10-156-761-3714
; Sequence 3714, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272637
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3714
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1137)
US-10-156-761-3714

Query Match          5.6%; Score 65.8; DB 15; Length 1137;
Best Local Similarity 48.9%; Pred. No. 1.3e-07;
Matches 280; Conservative 0; Mismatches 272; Indels 21; Gaps 3

```

```

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHAYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OR INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
Query Match 5.6%; Score 65.8; DB 15; Length 9025608

```


Mon Apr 18 09:47:30 2005

US-09-855-340b-1.rnpb

Pa

665 TCGCCGCGCGGCCCCGCTGACCGTCTGACAGCTCCAGAGCTGGCGACGCGGAG 724
676 --ACCGAACACACCATCAACCGTCCGCAAGGCCAGGAGTGCACCGCTGACGACGAGGAG 733
725 AGCTGCTTTCCAGTCCGCGAAGACCGCGAAGGGCGCGGCGACGCTTCCACACG 784
734 AGCGCTTTGACAAAGATCCCAAGTCCGAGGCGGCGAAGGACCATCGGCTTCCGCTTCC 793
785 AAGTGCCTTACTGCTTACGCACTGCGCGGCGAAGAAAGTGAAGGCTGCTTCA 844
794 AGCTGCTTACCGTGTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 853
845 CCGCGCGCGAAGCGGAGTGTAAAGACCGCAATTTCCGCGCGATCTGGCTCAAGCGT 904
854 TAGCGCGACGAGTACCCCTGCGGCGGACACCTGTACAGAGGCTTGTAAAGCGCTC 913
905 GCGAGAGACCGCGGCTTCCGCGGCTTACGCACTTACGATCTGCGGCGACACTCAGCGGCG 964
914 GGGGAAAGTCCGACTGACACCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 973
965 TCTGATTTCTGCGCGGCGCTTCCGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1024
974 TCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1033
1025 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1084
1034 TGGCTG---CGGCTGCGCGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1090
1085 CCGCGATCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1111
1091 AAGCCCTTTCGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1117

RESULT 7
US-09-821-167-1
Sequence 1, Application US/09821167
Patent No. US20020015989A1
GENERAL INFORMATION:
APPLICANT: Hosted Jr., Thomas J.
APPLICANT: Horan, Ann C.
TITLE OF INVENTION: Isolated Nucleic Acids from *Micromonospora rosaria*
FILE REFERENCE: IN011490
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: US 60/194,461
PRIOR FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1118
TYPE: DNA
ORGANISM: *Micromonospora rosaria*
FEATURE:
NAME/KEY: RBS
LOCATION: (6055)..(6059)
NAME/KEY: RBS
LOCATION: (6391)..(6394)
NAME/KEY: RBS
LOCATION: (8084)..(8088)
NAME/KEY: RBS
LOCATION: (9834)..(9837)
NAME/KEY: RBS
LOCATION: (10010)..(10012)
US-09-821-167-1

Query Match
Best Local Similarity 4.8%; Score 56.6; DB 9; Length 1118;
Matches 262; Conservative 0; Pmed. No. 3.6e-05; Indels 6; Gaps 2;

545 AGATCGGCTTATACGCGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 604

10584 AGTGTGTCCTGCTGAGGCGCATCAACCGCGCTTACCGGACGCGGCTCTGATTCGCG 10584
605 TGGGACCGGCTGTAGGTGGGCTAGAGGCGATCGGCTTCCGCGGCGGCGGCGGCGGCGGCT 605
10644 CCGTGTGCGGCTTACCGCGGAGAGATGCGAGGACTCCGGGTCCGACAGGCTGAGCTG 10644
665 TCGCCGCGGCGGCGGCGGCTGACCGTCTGAGCAGCTCCAGGAGCTGCGCAGCAGGAG 665
10703 --ACCGAACACACCATCAACCGTCCGCAAGGCCAGGAGTGCACCGCTTCCGCTTCC 10703
725 AAGTGCCTTACTGCTTACGCACTGATGCGCGGCGAAGAAAGTGAAGGCTGCTTCA 725
10761 AAGCGCTTACAGAGATCCCAAGTCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 10761
785 AAGTGCCTTACTGCTTACGCACTGATGCGCGGCGAAGAAAGTGAAGGCTGCTTCA 785
10821 AGCTGTACCGGCTGATCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 10821
845 CCGCGCGCGAAGCGGAGTGTAAAGACCGCAATTTCCGCGGCGGATCTGGCTCAAGCG 845
10881 TCAGCGCGACGAGTACCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 10881
905 GCGAGAGACCGGCGGCTTCCGCGGCTTACGCACTTACGATCTGCGGCGGCGGCGGCGG 905
10941 GGGGAAAGTCCGACTGACACCTTCACTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCG 10941
965 TCTGATTTCTGCGCGGCGCTTCCGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 965
11001 TCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 11001
1025 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1025
11061 TGGCTG---CGGCTGCGCGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 11061
1085 CCGCGATCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1111
11118 AAGCCCTTTCGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 11144

RESULT 8
US-10-156-761-3695
Sequence 3695, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 3695
LENGTH: 1179
TYPE: DNA
ORGANISM: *Streptomyces avermitilis*
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1179)
US-10-156-761-3695

Query Match
Best Local Similarity 4.7%; Score 55.8; DB 15; Length 1179;
Matches 117; Conservative 0; Pmed. No. 5.7e-05; Indels 0; Gaps 0;

QY 837 CATTTCACCGCCGCGAAGGCGGAGATGTAAGACGCGCAATTTCCGGGAGATCTGGGT 896
 Db 858 CGTTTCGCGCGGAGACGCGCGCGGCTGACGCGCGGAGCGTGTACCGGCTTTTAC 917
 QY 897 CAAGGCGTGCAGAAAGCCGGGCTTCCGGCTTACGATTCAGGATCTGGGACACTCA 956
 Db 918 GCGGCTCGCGGCTGACGAGGCTGAGGGTGTCCGGCTGACGAGCGGAGCGGCTG 977
 QY 957 CGCGGAGATCTGATTTTCCGGGCGTCCGCTGTGCGGATCTCCCGCGCTCGGTCA 1016
 Db 978 CGCACGCTACTGACCGCGCGCGCTCCGCGCGGCTCATGTGAATCTCGGGCA 1037
 QY 1017 CTCGTCATCGCGGTACCGGATCTGCTGTAACGGGACCT 1055
 Db 1038 CAGCGAGATCAGATCAGATGAGCGTGTACGACGACGT 1076

RESULT 9
 US-10-203-295-2/c
 ; Sequence 2, Application US/10203295
 ; Publication No. US20040115762A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zolchev, Sergey Borisovich
 ; APPLICANT: Sekurova, Olga Nikolayivna
 ; APPLICANT: Fjaervik, Epsen
 ; APPLICANT: Brautaset, Trygve
 ; APPLICANT: Strom, Arne Reidar
 ; APPLICANT: Valla, Svein
 ; APPLICANT: Ellingsen, Trond Erling
 ; APPLICANT: Sletta, Havaid
 ; APPLICANT: Gulliksen, Ole-Martin
 ; TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their
 ; FILE REFERENCE: 1181-265
 ; CURRENT APPLICATION NUMBER: US/10/203,295
 ; PRIORITY FILING DATE: 2003-05-19
 ; PRIOR APPLICATION NUMBER: PCT/GB 01/00509
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: GB 0002840.7
 ; PRIOR FILING DATE: 2000-02-08
 ; PRIOR APPLICATION NUMBER: GB 0008786.6
 ; PRIOR FILING DATE: 2000-04-10
 ; PRIOR APPLICATION NUMBER: GB 0009387.2
 ; PRIOR FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 27541
 ; TYPE: DNA
 ; ORGANISM: Streptomyces noursei ATCC 11455
 US-10-203-295-2

Query Match 4.7%; Score 55.8; DB 18; Length 27541;
 Best Local Similarity 44.8%; Pred. No. 6e-05;
 Matches 346; Conservative 0; Mismatches 417; Indels 10; Gaps 3;
 QY 297 TGACAGCTGACGCGGACAGTCAACAGATGAGTCAACGATCGGAGCGCGGCTGG 356
 Db 4651 TGTGTGATGGGCGTGTGGCGCTGTGTAATGCCCGCGCTCGGTGGCGCGCT 4592
 QY 357 CCCGCGCGGAGTCCAGCGCGGCTGTGGAAGCCGTGGAGGAGAAAGATCAGCA 416
 Db 4591 GCGCGCGCGGAGCGGAGTGGCGCGATCTGGGTCGCGCTTCCGCTTCCGCGCAT 4532
 QY 417 CTGCGACGCGCTGCTGACAGATCTGCGCGCGCGGCGGAGGAAAGATCAGGCT 476
 Db 4531 CTGGAATCTTCCGCGCGGAGATCGCGGTCGCGGACCCCTCCGCTGACCGGTC 4472
 QY 477 CAACCCGCTCTTGAAGATCTGCGCGCGGCGGAGCGGAGAAAGATGAACTTCTGAG 536
 Db 4471 GGTGAAGAGCTGACGAGGTGCAAACTTCGCCAAGCGGCTTCGGACTTCGTCTG 4412
 QY 537 CGACCCGAGATCGGTGCTTATCAGGCGCTTCGCGCGGACTG--CGACCGCTCGT 593

Db 4411 CGCGCGCTGATGTCCTTGGGACGCTGTGCTGCGCTGCGCAGAGCTTCCGCTGGC 4352
 QY 594 CATGCTGCTGATGCGACCGGCTGTGAGTGGGGTGAAGGATCGGCTTCGCGCGCG 653
 Db 4351 ACTGCTGCTGATGCGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4292
 QY 654 GGTGACCTGCTGCGCGCGCGCGCGCGCGCTGACCGCTGTC-----GAGCAGCTCAGGA 707
 Db 4291 GATGGGACCCCTGTACGCGCGGATCAGTACCTTGACCGGCTGGGGCGGCTCTGG 4232
 QY 708 GCTGCGAGGACCGGAGAGCTGTCTTCAAGTGCAGGAGACCGGAGAGGCGCGGAC 767
 Db 4231 CGAGGCGATCACCGGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4172
 QY 768 GATCAGTTTACCAAGAAAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 827
 Db 4171 GCGGTGCGCGAGACCAACGAGCGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4112
 QY 828 TGACGAGTCTGTGTACCGCGCGGAGGAGGAGGATGATGAGACCGGCAATTTCCGCG 887
 Db 4111 CGCGAGATGCTGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4052
 QY 888 GATTCGCTCAAGGCGGTGCGAGAGCGCGGCTTCCGCGCTTACGATTCAGATCTGCG 947
 Db 4051 GACCGGCTCCACCGGATCGACGCGGCGGAGGCGGATGCGCTGCTGCTGCTGCTG 3992
 QY 948 GCACTACAGCGCGGATCTGTATTGTCG--GCGCTGCTGTGCGGATCTCCGCGC 1006
 Db 3991 GAGCTATCTCTGCTGTATCTGTATGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3932
 QY 1007 GCTCGGTGATCTGTGATCGCGGTACGAGATCTGCTGTACGCGGACCTGCGT 1059
 Db 3931 GCGCGGCGCGGCTGTGTGCGCGGATCGCGGAGTCTGTGACACCGCGGT 3879

RESULT 10
 US-10-203-295-35/c
 ; Sequence 35, Application US/10203295
 ; Publication No. US20040115762A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zolchev, Sergey Borisovich
 ; APPLICANT: Sekurova, Olga Nikolayivna
 ; APPLICANT: Fjaervik, Epsen
 ; APPLICANT: Brautaset, Trygve
 ; APPLICANT: Strom, Arne Reidar
 ; APPLICANT: Valla, Svein
 ; APPLICANT: Ellingsen, Trond Erling
 ; APPLICANT: Sletta, Havaid
 ; APPLICANT: Gulliksen, Ole-Martin
 ; TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their
 ; FILE REFERENCE: 1181-265
 ; CURRENT APPLICATION NUMBER: US/10/203,295
 ; PRIORITY FILING DATE: 2003-05-19
 ; PRIOR APPLICATION NUMBER: PCT/GB 01/00509
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: GB 0002840.7
 ; PRIOR FILING DATE: 2000-02-08
 ; PRIOR APPLICATION NUMBER: GB 0008786.6
 ; PRIOR FILING DATE: 2000-04-10
 ; PRIOR APPLICATION NUMBER: GB 0009387.2
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 35
 ; LENGTH: 125401
 ; TYPE: DNA
 ; ORGANISM: Streptomyces noursei ATCC 11455
 US-10-203-295-35

Query Match 4.7%; Score 55.8; DB 18; Length 125401;
 Best Local Similarity 44.8%; Pred. No. 6.1e-05;

Mon Apr 18 09:47:30 2005

us-09-855-340b-1.rmpb

Pa

Matches 346; Conservative 0; Mismatches 417; Indels 10; Gaps 3;

QY 297 TGACGAGCTGACGCGGAGGTACCCAGCAGTGGGTCAACGACTGAGGCGCGGTGG 356

Db 4651 TGTGTGATGAGGCGGTGTGGCTGTGTGATGCGCCCGCGGTGTGGCGCGCGCT 4592

QY 357 CCCGTGCGCGAGTTCACGCGGCGGTGTGTGAGAGCGCTGTGGAGGAGAGATGAGCA 416

Db 4591 GCGCGCGCGGACCGGACGCGGCGGTGTGTGAGAGCGCTGTGTGTGTGTGTGTGT 4532

QY 417 CTGCGACGCGCTGT 476

Db 4531 CCTGACCTTCTTCCGCGCGGAGATGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 4472

QY 477 CAACCCGCTGT 536

Db 4471 GGTGAACGACGTGACGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4412

QY 537 CGACCGGAGATCGGT 593

Db 4411 CGCGCGCGGT 4352

QY 594 CATGCTGT 653

Db 4351 ACTGCTGT 4292

QY 654 GGTGACCTGT 707

Db 4291 GATGGGACCTGT 4232

QY 708 GCTGGGACGACGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 767

Db 4231 CGAGGCGATCACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4172

QY 768 GGTGACCTGT 827

Db 4171 GCGGTTCGCGGACGACGACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4112

QY 828 TGACGAGT 887

Db 4111 CGGACGAGT 4052

QY 888 GATGTGGGTCAAGCGGT 947

Db 4051 GACCGGT 3992

QY 948 GCACTGT 1006

Db 3991 GAGCTATCTGT 3932

QY 1007 GCCTCGGTCACTGT 1059

Db 3931 GCCGCGGCGCGGT 3879

RESULT 11

US-10-260-5339/c

Sequence 5339, Application US/10260238

Publication No. US20040016025A1

GENERAL INFORMATION:

APPLICANT: Budworth, Paul R.

APPLICANT: Moughamer, Todd G.

APPLICANT: Briggs, Steven P.

APPLICANT: Cooper, Bret

APPLICANT: Glazebrook, Jane

APPLICANT: Goff, Stephen A.

APPLICANT: Katagiri, Fumiyaki

APPLICANT: Kreps, Joel

APPLICANT: Provatt, Nicholas

APPLICANT: Rieke, Darrell

CURRENT APPLICATION NUMBER: US/10/260,238

CURRENT FILING DATE: 2002-09-26

PRIOR APPLICATION NUMBER: US 60/325,448

PRIOR FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: US 60/325,277

PRIOR FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: US 60/370,620

PRIOR FILING DATE: 2002-04-04

NUMBER OF SEQ ID NOS: 6077

SEQ ID NO 5339

LENGTH: 846

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc_feature

LOCATION: (345)..(345)

OTHER INFORMATION: n = any nucleotide

US-10-260-5339

Query Match 4.6%; Score 54.4; DB 17; Length 846;

Best Local Similarity 46.6%; Pred. No. 0.00013; Indels 0; Gaps 0;

Matches 172; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 369 GTCCACGCGGCGGT 480

Db 480 GGTGACGCGGAGCGGATGT 429

QY 429 GGTGACGAGT 420

Db 420 GCGCGGCGGT 489

QY 489 TTGACGATGT 360

Db 360 CATGGCGGT 549

QY 549 CGGT 300

Db 300 CCGCGCTGT 609

QY 609 GACCGT 240

Db 240 GACGAGT 669

QY 669 CGCGCGGCGCGGT 180

Db 180 CGCGCGGCGCGGT 729

QY 729 CGTCTTCA 737

Db 120 CGAGCGCGA 112

RESULT 12

US-10-156-761-5080

Sequence 5080, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OR INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

SEQ ID NO 5080
LENGTH: 879
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(879)
US-10-156-761-5080

Query Match 4.5%; Score 53.6; DB 15; Length 879;
Best Local Similarity 44.0%; Pred. No. 0.00022;
Matches 227; Conservative 0; Mismatches 289; Indels 0; Gaps 0;

229 ACCGCGTGAAGTCCGAGGCAACCGGATCCGCAACCTCTCCGCAATCTCGGCAAT 288
274 ACCGCTGATGATCAGGAGCGCGGAGACCTCTCTGTCTGTCTATCGCGGATGCGCGC 333
289 CTACACCTTGACGAGCTGAGCGGAGGTCACCCAGACAGTGGTCAACGACTGAGAGC 348
334 TACGCTTCGCGTGAAGTCACTTCAAGGGCCGCGATGGTGTTCATGGTGTGTGGCG 393
349 GGGTGGCGCCGCGCGGAGTTCACCGGGGCTGTGGAAGCCGTGGCAGGAAGAG 408
394 CTGCTGTGTGTGCGGAGTGAAGTGGCGCTGATCCGCTCTCGATCTTTCGCAAGTTC 453
409 ATCAGCACTGCCACGCGCTGTGCAACAGATCTGGCGCGCGCATTCGCGCAAGAG 468
454 GGGATCTTGGGAGCATTCATCGCGCTGTCTTTCATGTGCGCTTCGCGCTCCGCTTC 513
469 ATCAGGCTCAACCGCTGTCTTTCAGATCTGCCCGCGCGAGCGCAAGAGATGAAG 528
514 GCGATCTTCTGTGTGGCAACTTTCGCGGAGATTCGCGGAGAGTGTGGAAGGCGCG 573
529 TTCTGAGGAGCCCGAGATCGGTGGCTTATCAGCGCGCTTCGCGGAGTGGCGAGCG 588
574 CGTCTGACGCGGCGGAGGCAAGGCGCTTTCGAGACCGTGTGTGCTGCGTGGCGCG 633
589 CTGCTGATGCTGTGTGTGCGGAGCGGCTGTGAGTGGGTGAGCGCATCGCGCTGCGCG 648
634 CCGGCAATCGGCTCCCTGGGAGTCTTTCAGATCTGTGGGTGTGAAGAGATGTGTGTC 693
649 GCGCGGCTGCACTGTGTGCGCGCGCGCGCGCTGAGCCGTGTGTGAGAGCTCCAGAG 708
694 GCGCTGTCTTCTGAGACTTCAGCTCGAGCGCGCTGAGCGGCTCGCTCCAGCAGAGGTG 753
709 CTGGCAGCAGCGGAGAGCTGTCTTTCAGTGGCGCG 744
754 CGGCACTTCCGCAACATGAGATCTGCGCGCG 789

RESULT 13
US-10-762-107-45
Sequence 45, Application US/10762107
Publication No. US20050043297A1
GENERAL INFORMATION:
APPLICANT: ECOPIA BIOSCIENCES INC.
APPLICANT: Farnet, Chris
APPLICANT: McAlpine, James
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Pirae, Mahmood
TITLE OF INVENTION: FARNESYL DIBENZODIAZEPINONE, PROCESSES FOR ITS PRODUCTION AND ITS
FILE REFERENCE: 3005-5US
CURRENT APPLICATION NUMBER: US/10/762,107
PRIOR FILING DATE: 2004-01-20
PRIOR APPLICATION NUMBER: USSN 60/441,126
PRIOR FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: USSN 60/492,997
PRIOR FILING DATE: 2003-08-07
PRIOR APPLICATION NUMBER: USSN 60/518,286
PRIOR FILING DATE: 2003-11-10
NUMBER OF SEQ ID NOS: 89

SOFTWARE: PatentIn version 3.0
SEQ ID NO 45
LENGTH: 1104
TYPE: DNA
ORGANISM: Micromonospora sp. strain 046-EC011
US-10-762-107-45

Query Match 4.5%; Score 53.2; DB 19; Length 1104;
Best Local Similarity 44.0%; Pred. No. 0.00028;
Matches 226; Conservative 0; Mismatches 288; Indels 0; Gaps 0;

91 ACCAGGCGCAAGATCCGATGATGATGCTCCGTGCGGAGACATTTGAGAGGCAACGCGCTC 150
361 ACCAGATCCCGCTGAGATGCTGTGGCGCGCGCGCGCGCGCGCGAGCATCATC 420
151 ATGCGCGCGCGCGGTCAATTAACCTCCGCGCATTTCTGTGGGAGATGTGTGCGAGCTAC 210
421 ACCGAGTGTCCGAGCTGAGAGGCGGAGATGTTCTTGTGGGTCTGTGAGCAGCGCTCG 480
211 GAAAGACGCTGAAACCGACCGCGTGAACCTGAGAGGCAACCGATCCGCAACCACTTC 270
481 GACGAGTGTATCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
271 CTGCCATATCTGCGCATTCACCTTGAACAGCTGAGCAGGCGAGTCAACCGAGTGG 330
541 GTGACACGCGCGCGAGCTGTGCTGTGAGTGTGAGGTACCGGATCCGCGCGCGG 600
331 GTCAACGACCTGAGAGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 390
601 GCGATGGCGAGCGCGCGCTGTGCGACAGTGAACGAACTTCCGTGTGAGCAGAGGCGATC 660
391 CCGCTGAGCGAAGAGATGATGAGCACTGCAACGCGCGCTGTGCAACGATCTGCGCGCG 450
661 CTGTGCGCTGTGCACTTCAACGCGAGTATCTGTGTGCAACGAGACCGATCCGCTGCG 720
451 GCGATGCGCGGAAACGAGTCAAGCTCAACCGGTCTTTCAGAGATGCTGCCCGCGCG 510
721 TACATGCGGACCGCGCGCTTCCGAGTCAACCGCGCGCGCTGCACTGTACACGCTTCC 780
511 GAGCGGAAAGATGAAGTCTTCTGAGGACCGCGAGATCGTGGCTTATCAGCGCGCT 570
781 GCGCGCGCGCGGACCACTACTACTGAGGAGTGTGTCTCCGCGCGCGCGCGCGCGCG 840
571 CCGCGCACTGCGGACCGCTGTCTATGCTGTG 604
841 GACTGCGAGGGAAGTCCCGGTGTGTACAGTGG 874

RESULT 14
US-10-762-107-1
Sequence 1, Application US/10762107
Publication No. US20050043297A1
GENERAL INFORMATION:
APPLICANT: ECOPIA BIOSCIENCES INC.
APPLICANT: Farnet, Chris
APPLICANT: McAlpine, James
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Pirae, Mahmood
TITLE OF INVENTION: FARNESYL DIBENZODIAZEPINONE, PROCESSES FOR ITS PRODUCTION AND ITS
FILE REFERENCE: 3005-5US
CURRENT APPLICATION NUMBER: US/10/762,107
PRIOR FILING DATE: 2004-01-20
PRIOR APPLICATION NUMBER: USSN 60/441,126
PRIOR FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: USSN 60/492,997
PRIOR FILING DATE: 2003-08-07
PRIOR APPLICATION NUMBER: USSN 60/518,286
PRIOR FILING DATE: 2003-11-10
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1

Mon Apr 18 09:47:30 2005

us-09-855-340b-1.rmpb

LENGTH: 36602
TYPE: DNA
ORGANISM: Micromonospora sp. strain 046-EC011
US-10-762-107-1

Query Match 4.5%; Score 53.2; DB 19; Length 36602;
Best Local Similarity 44.0%; Pred. No. 0.0029;
Matches 226; Conservative 0; Mismatches 288; Indels 0; Gaps 0;

91 ACCAGCCCAAGATGCGATGTCGATTCGCTGGCGAGCAGTTCAGAGGCAACCGCTC 150
27395 ACCAAGATCCCTCGAGATCTGCTGGCGCGCGCGCGCGCGCGAGGCAAGCATATC 27454
151 ATGCCCG 27514
27455 ACCCAGGTGCG 27514
211 GAAAGACGCTGAAACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 270
27515 GACGAGTATCTGCG 27574
271 CTGCCATATCTGCG 330
27575 GTGAGCAGCG 27634
331 GTCAACGACTGAGCG 390
27635 GGCATGGCG 27694
391 CGCTGGCAGCG 450
27695 CTGGTCGCTGCG 27754
451 GCGATGCG 510
27755 TACATGCG 27814
511 GAGCGCAAGAGATGAGATTCCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 570
27815 GCCG 27874
571 CCG 604
27875 GACTCGCAGCG 27908

RESULT 15

US-10-156-761-5794
Sequence 5794, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ. ID NOS: 15109
SEQ. ID NO 5794
LENGTH: 426
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS

LOCATION: (1)..(426)
US-10-156-761-5794

Query Match 4.4%; Score 52.2; DB 15; Length 426;
Best Local Similarity 52.0%; Pred. No. 0.0005;
Matches 142; Conservative 0; Mismatches 128; Indels 3; Gaps 0;

831 CGAGGCTGTGTTACCG 426
117 CGGGCTGCTTCTGCG 426
891 CTGGGTCAAGCGGTGCG 426
177 CTGGCTCGAGGCGCTGACG 426
951 CACTCAGCG 426
237 CACCGGGAACAGCTGCG 426
1011 CGGTGACTGCG 426
297 GGGCCACTCTCTGCG 426
1071 CGAGGCGATCTGCG 1103
354 CACACAGATCG 386

Search completed: April 15, 2005, 22:50:53
Job time : 891.841 secs

THIS PAGE BLANK (USPTO)

Mon Apr 18 09:47:29 2005

US-09-855-340b-1.rn1

Pa

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 17:52:06 ; Search time 249.611 Seconds
(without alignments)
7728.636 Million cell updates/sec

Title: US-09-855-340b-1

Perfect score: 1179
Sequence: 1 gcttgatcgcagagaagacg.....tgctgacgcagcagcatga 1179

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgnt2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgnt2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgnt2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgnt2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgnt2_6/ptodata/1/ina/PTCUTS.COMB.seq.*
6: /cgnt2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61.6	5.2	2404	4	Sequence 16, Appl
2	57.2	4.9	1806	4	Sequence 3884, Ap
3	57.2	4.9	21143	4	Sequence 1191, Ap
4	57	4.8	4719	4	Sequence 5392, Ap
5	57	4.8	11282	4	Sequence 1039, Ap
6	56.6	4.8	1161	4	Sequence 15, Appl
7	56.6	4.8	11188	4	Sequence 3606, Ap
8	53.8	4.6	1329	4	Sequence 1173, Ap
9	53.8	4.6	20113	4	Sequence 6567, Ap
10	52.8	4.5	879	4	Sequence 490, App
11	52.8	4.5	2760	4	Sequence 1191, Ap
12	51.4	4.4	516	3	Sequence 34, Appl
13	51.4	4.4	516	3	Sequence 34, Appl
14	51.4	4.4	516	3	Sequence 34, Appl
15	51.4	4.4	516	3	Sequence 34, Appl
16	51.4	4.4	516	3	Sequence 34, Appl
17	51.4	4.4	516	3	Sequence 34, Appl
18	51.4	4.4	516	3	Sequence 34, Appl
19	51.4	4.4	516	3	Sequence 34, Appl
20	51.4	4.4	516	3	Sequence 34, Appl
21	50.2	4.3	534	4	Sequence 11386, A
22	50.2	4.3	3150	4	Sequence 11487, A
23	50.2	4.3	3372	4	Sequence 11405, A
24	49.8	4.2	887	4	Sequence 8464, Ap
25	49.8	4.2	1407	4	Sequence 5975, Ap
26	49.8	4.2	1513	4	Sequence 210, App
27	49.8	4.2	1674	4	Sequence 5626, App

28	49.8	4.2	7012	4	US-09-902-540-890	Sequence 890.
29	49.8	4.2	43225	4	US-09-902-540-1269	Sequence 1269
30	49.4	4.2	3244	4	US-09-902-540-9009	Sequence 9009
31	49.4	4.2	12849	4	US-09-902-540-963	Sequence 963
32	49.2	4.2	1473	4	US-09-902-540-2975	Sequence 2975
33	49.2	4.2	2415	4	US-09-902-540-354	Sequence 354
34	48.8	4.1	915	4	US-09-252-991A-2186	Sequence 2186
35	48.8	4.1	1036	4	US-09-252-991A-2345	Sequence 2345
36	48.8	4.1	936	4	US-09-252-991A-2490	Sequence 2490
37	48.6	4.1	1059	4	US-09-252-991A-15538	Sequence 15538
38	48.6	4.1	744	4	US-09-252-991A-15427	Sequence 15427
39	48.6	4.1	807	4	US-09-724-797-39	Sequence 39
40	48.2	4.1	1137	4	US-09-902-540-7537	Sequence 7537
41	48.2	4.1	1467	4	US-09-902-540-1058	Sequence 1058
42	47.8	4.1	1467	4	US-09-902-540-2886	Sequence 2886
43	47.8	4.1	762	4	US-09-902-540-9461	Sequence 946
44	47.8	4.1	2007	4	US-09-902-540-1077	Sequence 1077
45	47.8	4.1	15095	4	US-09-902-540-1139	Sequence 113

ALIGNMENTS

RESULT 1
US-09-269-717-16
Sequence 16, Application US/09269717
Patent No. 683135
GENERAL INFORMATION:
APPLICANT: FRAZAO MONIZ PEREIRA, Jose A.
APPLICANT: FREITAS VIEIRA, Alcinio
APPLICANT: RIBEIRO DOS SANTOS-JANES, Elsa M.
APPLICANT: DA COSTA GARCIA, Miguel A.
APPLICANT: DA SILVA ALVES, Paulo J.
TITLE OF INVENTION: DNA INTEGRATION INTO "MYCOBACTERIUM SP." GENOME
TITLE OF INVENTION: TRANS-COMPLEMENTATION USING A SITE-SPECIFIC INTE
FILE REFERENCE: 1780-104P
CURRENT APPLICATION NUMBER: US/09/269, 717
CURRENT FILING DATE: 1999-06-03
PRIOR APPLICATION NUMBER: PCT/PT97/00005
PRIOR FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 2404
TYPE: DNA
ORGANISM: Mycobacterium phage Ms6
FEATURE:
NAME/KEY: CDS
LOCATION: (697)..(1812)
US-09-269-717-16
Query Match 5.2% Score 61.6; DB 4; Length 2404;
Best Local Similarity 48.5%; Pred. No. 0.00035;
Matches 391; Conservative 0; Mismatches 364; Indels 51; Gap

317 TCACCAAGAGTGGGTCAAGACTGAGAGCCGGGCTGGCCCGGAGTCCA
1055 ACATGCGCCCGGTGGTGAAGGTCATGGAACCAACCGGCGG-----CGACGGCA
377 GGGGTGTCGAGAGCCGCTGGAGGAGAGATGAGCACTGCCAGGCTGCTGC
1110 GCACGCGCCCGGAACCTTCGGAACAAATACG---GTTCATATGGGGGACTGAA
437 CGATCGCGCGCGGAGATCGCGGGAACGATCAAGCTCAACCGTCTTGA
1167 CGCGGTCGCGGATCTTGTCAACCAACCTCGTGGCGCGCGCTGCGTGGG
497 TGCTGCGCGCGCGGAGCGGCAAGAGATGAAGTTCCTGAGCGACCGGAGATCGTC

Db 7000 CGGGGCGAGCGGCCCCGCGCGCGGTGAGAGCGCCCTTGCCCCCGAGAGAGTCCGCGC 6941
Qy 418 TGCCACGCGCTGCTGACACAGATCTGCGGCGCGGATGCGCGGAGAAAGATCAAGGCTC 477
Db 6940 GAGCGCGCCAGACCTGCGCGCTCCAGCGCTGCGGCTGCGCGGAGAGAGTTCGCGCGC 6881
Qy 478 AACCGTGCTTTCAGCATGCTGCCCCGCGCGGAGCCGAAAGATGAAGTTCTGAGC 537
Db 6880 GAACGTGAGACAGACAGAGAGTGTCTGCGGCGCTTCGGAAGTGTGATGAGCGCTTC 6821
Qy 538 GACCGGAGATGCTGCTTATACAGCGGCTTCGCGCGCATCTGCGGAGACCGCTCTCATG 597
Db 6820 GCGCTGACTCCATGATGAGCGGCGCGCGCGCGCGCATCTGCGCGGCTTGAACCGC 6761
Qy 598 CTGCTGCTGCGCGCGCGCTGAGGTGAGGCGCATCGCGCGCGCGCGCGCGCGGCTC 657
Db 6760 GTGCGGCTGCGCATGAGCGAGCTCTACGCACTGAGCGCGCATCCCGCGCGCTTACAGACG 6701
Qy 658 GACCTGCTGCG 717
Db 6700 ACCGCGCGCGCGCTGCTGCGCGCATGTAAGGCGCGCGCGCTTCAGACAGAGCTCGAGCGC 6641
Qy 718 ACGGAGAGCTGCTTCCAGTCCGCGAGAGACCGCGAGAGCGCGCGCG 764
Db 6640 CTGGGACGCTTGAGCGTCTTACGCGCATATGACCCGCGCGCGCTGCGC 6594

RESULT 4
US-09-902-540-5392
; Sequence 5392, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5392
; LENGTH: 4719
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(4719)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-5392

Query Match 4.8%; Score 57; DB 4; Length 4719;
Best Local Similarity 47.6%; Pred. No. 0.0039;
Matches 165; Conservative 0; Mismatches 182; Indels 0; Gaps 0;
Qy 398 CAGCGAAGCATGACGACGCTGCTGACACAGATCTGCGCGCGCGCATG 457
Db 3929 CATTGACATACCCGCGCGCGCGCGCGCATGCTGCTGACAGCCGACCTGT 3988
Qy 458 CGCGAAGCATGACGCTCAACCGCTGCTTTCAGCATGCTGCGCGCGCGCGCGCG 517
Db 3989 CCACTTTCGACCGGAGACCAACGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTG 4048
Qy 518 AAGAGATGAAGTTCTTACAGGACCGGAGATCGGTGCTTATACAGCGCGCTTCGCGCGC 577
Db 4049 AAGGATGACAGGAGATGAAGCGCGCGCGGAGGAGGCGGTGAAGATTG 4108
Qy 578 ACTGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 637
Db 4109 CCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4168

Qy 638 GCTGCGCGCGCGCGCGCGCTGACACCTGCTGCGCGCGCGCGCGCGCGCGCTGACAG 4275
Db 4169 GCGGCTTACAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4275
Qy 698 AGCTCCAGAGAGCTGCGCGAGAGAGCTGCTTCAAGTCCGCG 744
Db 4229 AGCCGAGCGGAGGAGGAGTTCATTCACAGCGCGCGCTTACTACTGCTG 4275

RESULT 5
US-09-902-540-1039
; Sequence 1039, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1039
; LENGTH: 11282
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(11282)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1039

Query Match 4.8%; Score 57; DB 4; Length 11282;
Best Local Similarity 47.6%; Pred. No. 0.0045;
Matches 165; Conservative 0; Mismatches 182; Indels 0; Gaps 0;
Qy 398 CAGCGAAGCATGACGACGCTGCTGACACAGATCTGCGCGCGCGCATG 457
Db 6803 CATTGACATACCCGCGCGCGCGCGCGCATGCTGCTGACAGCGCGGACT 457
Qy 458 CGCGAAGCATGACGCTCAACCGCTGCTTTCAGCATGCTGCGCGCGCGCGCGCGCG 517
Db 6863 CCACTTTCGACCGGAGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4048
Qy 518 AAGAGATGAAGTTCTTACAGGACCGGAGATCGGTGCTTATACAGCGCGCTTCGCGCGC 577
Db 6923 AAGCATGAGGAGGAGATGAAGCGCGCGCGGAGGAGGCGGTGAAGATT 4108
Qy 578 ACTGCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 637
Db 6983 CCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4168
Qy 638 GCTGCGCGCGCGCGCGCGCTGACACCTGCTGCGCGCGCGCGCGCGCGCGCTGACAG 4275
Db 7043 GCGGCTTTCGACCGGAGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4048
Qy 698 AGCTCCAGAGAGCTGCGCGAGAGAGCTGCTTCAAGTCCGCG 744
Db 7103 AGCCGAGCGGAGGAGGAGTTCATTCACAGCGCGCGCTTACTACTGCTG 4275

FILE OF INVENTION: plasmid pMR2 and Vectors Made Therefrom
FILE REFERENCE: IN011490
CURRENT APPLICATION NUMBER: US/09/821,167
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: US 60/194,461
PRIOR FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 1161
TYPE: DNA
ORGANISM: Micromonospora rosaria
US-09-821-167-15

Query Match
Best Local Similarity 46.2%; Score 56.6; DB 4; Length 1161;
Matches 262; Conservative 0; Mismatches 299; Indels 6; Gaps 2;

545 AGATCGGTGCGCTTATACAGCGGCTTCGCGGACCTGCGGACCGCTCGTCACTGCTGCG 604
557 AGGTGTCGCGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 616
605 TGGCGACCGGTCTGAGGTGCGGAGGATGCGGACCTCGGCTCGGACCGCTGCG 664
617 CCGTGGCGGCTTACCGCGGAGAGATGCGGACCTCGGCTCGGACCGCTGCG 675
665 TCGCGCGCGCGCGCGCTGACCGCTGCGGACCGCTGCGGACCGCTGCGGAG 724
676 --ACCGAACACACATCAACCGCTGCGGACCGCTGCGGACCGCTGCGGAG 733
725 AGCTGCTTTCAGTCCGCGGACCGCTGCGGACCGCTGCGGACCGCTGCGGAG 784
734 AGCGGTTCAGAGATCCCAAGTCCGAGCGGCGCAACCAATGCGCTCCCTCCC 793
785 AAGTCTCTACTGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 844
794 AGGTGTCGCGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 853
845 CCGCGCGCGCGCGCGCTGAGGACCGGCAATTCGCGGACCGCTGCGGAG 904
854 TCGCGCGCGAGATCAACCGCTGCGGCGGACCGCTGTAACAGCGCTGCGGAG 913
905 GCGAGAAAGCGGCGCTTCGCGGCTTACAGTTCAGTTCGCGGACCGCTGCGGAG 964
914 GGGGAAAGTTCGAGCTGCAACCGCTTCAACGACCTGCGGACCGCTGCGGAG 973
965 TCGTATTTTTCGCGGCGCTTCGCGGCTTCGCGGACCGCTGCGGAGCTGCGGAG 1024
974 TCGCGCGCGAGACCGGCGGCGACCTGCGGCGGACCTGTAAGCGGCTGCGGAGCTGCGGAG 1033
1025 TCGCGGTACAGGATGCTGTAAGGAGCACTGCGTGAAGGAGTTCGAGAGGAGATCTCG 1084
1034 TGGCTG---CGGCTGCGCGGATCTTCAAGCGCTGAGCGGTGACGCGGAGATGCGCA 1090
1085 CCGCGATTCGAGAGCGGATGCGGCGG 1111
1091 AAGCCCTTTCGAGCTGCGGCGGACG 1117

RESULT 7
US-09-821-167-1
Sequence 1, Application US/09821167
Patent No. 6569668
GENERAL INFORMATION:
APPLICANT: Hosted Jr., Thomas J.
APPLICANT: Horan, Ann C.
TITLE OF INVENTION: Isolated Nucleic Acids from Micromonospora rosaria
FILE REFERENCE: IN011490
CURRENT APPLICATION NUMBER: US/09/821,167
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: US 60/194,461
PRIOR FILING DATE: 2000-04-04

NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1118
TYPE: DNA
ORGANISM: Micromonospora rosaria
FEATURE:
NAME/KEY: RBS
LOCATION: (6055)..(6059)
NAME/KEY: RBS
LOCATION: (6391)..(6394)
NAME/KEY: RBS
LOCATION: (8084)..(8088)
NAME/KEY: RBS
LOCATION: (9834)..(9837)
NAME/KEY: RBS
LOCATION: (10010)..(10012)
US-09-821-167-1

Query Match
Best Local Similarity 46.2%; Score 56.6; DB 4; Length 1118;
Matches 262; Conservative 0; Mismatches 299; Indels 6; Gaps 2;

545 AGATCGGTGCGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 604
10864 AGGTGTCGCGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 10643
605 TGGCGACCGGTCTGAGGTGCGGAGGATGCGGACCTCGGCTCGGACCGCTGCG 664
10644 CCGTGGCGGCTTACCGCGGAGAGATGCGGACCTCGGCTCGGACCGCTGCG 10702
665 TCGCGCGCGCGCGCGCTGACCGCTGCGGACCGCTGCGGACCGCTGCGGAG 724
10703 --ACCGAACACACATCAACCGCTGCGGACCGCTGCGGACCGCTGCGGAG 10760
725 AGCTGCTTTCAGTCCGCGGACCGCTGCGGACCGCTGCGGACCGCTGCGGAG 784
10761 AGCGGTTCAGAGATCCCAAGTCCGAGCGGCGCAACCAATGCGCTCCCTCCC 10820
785 AAGTCTCTACTGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 844
10821 AGGTGTCGCGCTTATACAGCGGCTTCGCGGACCGCTCGTCACTGCTGCG 10880
845 CCGCGCGCGCGCGCGCTGAGGACCGGCAATTCGCGGACCGCTGCGGAG 904
10881 TCGCGCGCGAGATCAACCGCTGCGGCGGACCGCTGTAACAGCGCTGCGGAG 10940
905 GCGAGAAAGCGGCGCTTCGCGGCTTACAGTTCAGTTCGCGGACCGCTGCGGAG 964
10941 GGGGAAAGTTCGAGCTGCAACCGCTTCAACGACCTGCGGACCGCTGCGGAG 11000
965 TCGTATTTTTCGCGGCGCTTCGCGGCTTCGCGGACCGCTGCGGAGCTGCGGAG 1024
11001 TCGCGCGCGAGACCGGCGGCGACCTGCGGCGGACCTGTAAGCGGCTGCGGAGCTGCGGAG 11060
1025 TCGCGGTACAGGATGCTGTAAGGAGCACTGCGTGAAGGAGTTCGAGAGGAGATCTCG 1084
11061 TGGCTG---CGGCTGCGCGGATCTTCAAGCGCTGAGCGGTGACGCGGAGATGCGCA 11117
1085 CCGCGATTCGAGAGCGGATGCGGCGG 1111
11118 AAGCCCTTTCGAGCTGCGGCGGACG 11144

RESULT 8
US-09-902-540-3606/c
Sequence 3606, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.

DB 6046 GGGACCGAAGCGGAAGCGGGCCACACAGACGGGTGAACACCCACTGTGGCCCGGGAAGAT 6085
 QY 456 CGC---GGCGAAACGGATCAGGCTCAACCCGTGCTCTTCAGCAGATGTGCCCCGGCGCGCA, 512

Dy 961 GCGATCCGATTCTGCCGGGCCTGTCGCGCATCCTCCGCGCGGTGACATC
||| ||| |
Db 688 GC---CCTGCTCACCGACTGGACCAAGCCGCTGGGCCAAGTGGGCAACGCCGCTGA

QY 1021 TCGATCGGGGACGAGATCTGTCGACGGGACCTGCGTGAAGAGGTGACAGAGGGATC 1080
 Db 745 GTCAATCGAGGCGGTGACATCTCCGGGCACCGCGCCGAGAGACTACCCGAGGGTACG 804
 QY 1081 CTCCTCGCGATCGAGAGGCGATGAGCCGGCTCCGGGCTGAGACTTGAGAGCGGAACTC 1140
 Db 805 TATGCCCTCCCGGCGGAGATCTGCTGCTGGGCAAGAACCCGCCCGGTGGAGAGGAG 864
 QY 1141 GACGAGGAGCTG 1152
 Db 865 GCGGAGGAGGTG 876

RESULT 11
 US-09-902-540-490/c
 Sequence 490. Application US/09902540
 Patent No. 6833417
 GENERAL INFORMATION:
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Wiegand, Roger C.
 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 FILE REFERENCE: 38-10(15849)B
 CURRENT APPLICATION NUMBER: US/09/902,540
 CURRENT FILING DATE: 2001-07-10
 PRIOR APPLICATION NUMBER: 60/217,883
 PRIOR FILING DATE: 2000-07-10
 NUMBER OF SEQ ID NOS: 16825
 SEQ ID NO 490
 LENGTH: 2760
 TYPE: DNA
 ORGANISM: Myxococcus xanthus.
 US-09-902-540-490

Query Match	Similarity	4.5%	Score 52.8	DB 4	Length 2760
Best Local	Similarity	44.7%	Pred. NO. 0.028		
Matches 247	Conservative	0	Mismatches 302	Indels 3	Gaps 1
QY	CTGTGGGAGCACCGGCTGTAGGTGGGGTATAGACGATCCGCGCTGGCGCCCGGGGTGCAC	601			660
Db	CTGGGGGACACAGGGGTGGGACCCCTGGACAAAGCTGAGTCCATCCCGGTTCAACGTCAAC	553			494
QY	CTGTCCGCGCGCGCGCGCGCGCTGACCGTGTGAGCAGCTCCAGACCTGGCCACAGCAG	661			720
Db	CTGCCAGAGTCCGAGATACCGCGCGGTGTGCGCGAGGTGAATCTGCTGCTATTGGCCAG	493			434
QY	GGAGAGCTCTCTTCCAGTCCGCCAGACACCGCGAAAGGGCCGGCGACGTCAGTTTACC	721			780
Db	ACGGCGCAGGTGGCCCCCGCGGACAAAGAGCTTATATGCCCTTGGCGGAGTGAACCGCGAC	433			374
QY	ACGAAAGTCGCTTACTGTACCGCACTCATCCGCGCGGAAAGAAATGACGAGGTCTGTG	781			840
Db	GTGAGATGCAATCCGCTATTGCGTCTTCCATCATGACAAAGAGCTGGCGGAGGCAATC	373			314
QY	TTCAACCGCGCCGAAAGGCGGAGTGTAAAGACGCGCAATTTCCGCGCGATCTGGTCTAAG	841			900
Db	GACGGCGTGTGTGCTGAGACGTGAAGGTGGGACGTGGCGCTTCATTAACGGGACGAGGAC	313			254
QY	GCGTGCAGAGAAAGCCGGGCTTCCGGGCTTACGACATTACAGATCTGCGGCACTACACGC	901			960
Db	GCGGCGACGCTGGCGGACAAACGATGATTTGGCTCTGGGCGCGGAGATGGGAAAGATGCTGTG	253			194
QY	GGGATTCGATTTTCTGCGCGGGGTCCGCGTGTCCGGGATCTCCCGCGGCTGGGTACTCG	961			1020
Db	GC---CTTGTCTCACGACATGACCTAGCCGCTGGGCCCGCAAGGTGGGCAACGCGCTGGAG	193			137
QY	TTCGATCCGCGTACCGGATCTGTGTACGGGACCTTCGCTGAGAGAGTGAAGGGGAGATC	1021			1080
Db	GTCAATTCAGGGGGGTGACATGTCTCCGGGCAACGCGCGGAGGACATCAACGAGGTGACG	136			77
QY	CTCGCGCGGATTCGAGAGGCGGATGCGCGGCTCGGCGCTGAGAGCTTGAAGCGGACATC	1081			1140

Db 76 TATGCCCTCCCCGCGAGATGCTGTGCTGGCGAAGAAGCCGCCCGGTGGAGAGGAG 17

QY 1141 GACGAGAGCTG 1152

Db 16 CGCGAAGAGTG 5

RESULT 12
 US-09-183-861-34
 ; Sequence 34, Application US/09183861
 ; Patent No. 6365165
 ; GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Campos-Neto, Antonio
 APPLICANT: Webb, John R.
 APPLICANT: Dillon, Davin C.
 APPLICANT: Skelky, Yasir A.W.
 TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND
 NUMBER OF SEQUENCES: 87
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 City: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/183,861
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/022,765
 FILING DATE: 12-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Makl, David J.
 REGISTRATION NUMBER: 31,392
 REFERENCE/DOCKET NUMBER: 210121.420G3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 516 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Leishmania major
 US-09-183-861-34

Query Match	4.4*	Score 51.4	DB 3	Length 516
Best Local Similarity	47.1*	Pred. No. 0.042		
Matches 222	Conservative 0	Mismatches 246	Indels 3	Gaps 2
Qy	302	AGCTGACACGGGACAGTCA	CGCCAGCAGTGGGTCA	CAGACTTGAGGCCGCTCGGCCGT 361
Db	7	AGCAGCCCTTAGACACCGC	CACGACGAGCGCGCGC	GAGCTGGAGGACCGGTGGCAGCGC 66
Qy	362	GGCCGAGTCCACGGGGGTG	TCGAGGCCGCTTGCAGCGA	-GACGATACGAACTGC 420
Db	67	TGCGCGCGACCGCGAGCG	CGCCACGAGCTGCGCGG	AGACGCCGAGCGTGCAGC 126
Qy	421	CACGGCCTGTGCACACGAT	CTGGCGCGCGGCATTCG	CGGGGAAACGATCAGCTCAAC 480
Db	127	AGCCCTTAGAACCGCACCG	ACGAGAGCGCGCCGAGCT	TGAGAGCACCGGTGGCACGCTGG 186

Db 7 AGCAGCCGCTTAGACACCGCCACGACGACGCGCCGAGCTGGAGGACACGGGTGGACACGG
Oy 362 GCGCGAGATCCACGCGGGGTGTCGAGAACCCGCTGCGAGCCGAA-GAGGATGAGCAACTG
Db 67 TGGCCGCGGAGACCGGAGACGAGCGCGCCAGCAGCTGGCCCGCAAGCCGAGGAGCTGCAG
Oy 421 CACGGGCTGTGCAACAGATTGGGGGGCGGGCCATCCGGGCGAAACGATCAGGCTCAA
Db 127 AGCGCTTAGACACCGCCACGACGACGCGCGCGGAGCTGGAGGACACGGGTGGACACGGCTG
Oy 481 CCGTGCTCTTCGACGATGCTGCCCCGCGCGCAGCGCCAAAGATGAAATTCTTAGCGCA
Db 187 CCGCGGACGGGAGACGAGGCCCGCCACGACGCTGGGCGCGGAACGCCAGGAGCTGCACGAG
Oy 541 CCGGAGATCGGTGGGCTTATACACGGCGCTTCGCGCCGACTGGGAGACGGCTGTATGCT
Db 247 GCTTAGACACCGCCACCGACGACGCGCCCGACGCTGAGGACACAGTGGGCAACGGCTGGCGC
Oy 601 CTGTGGCGCACCGGCTTAGGTGGGGTGAAGGCCAT--CGGCTGTCGCGCGCGCGCGGGTTC

302 AGCTGACGGGCAGGTCACCCAGCAGTGGGTCAACGACTGGAGGCCGGCGTGGCCCGT 361

481 CCGGCTCTGACGAGCTGCCCCGCGCG

QY 541 CCGGAGATGCTGCTTATCAAGGCGCTTCCGCGCACCTGGGACCGCTCATGCTG 600
 DB 247 GCTTATACACCCGACGACGAGCGCGAGCTGAGGACAGGTGGACAGCTG6CCG 306
 QY 601 CTGATGGCCAGCCGCTGTAGTGGGGTGAAGCGAT--CGGCTTCCGCGCGCGGCTG 658
 DB 307 CGAAGCCCGAGAGCTGCGACGAGCGCTTACACCGCCACGCGAGCGCGCGAGCTG 366
 QY 659 ACCTGCTCCGCGCGCGCGCGCGCTGACCGCTGTCAGACGCTTCAGAGAGCTGCCAGCA 718
 DB 367 AGGCAAGGGGTGGACCGGCTGCGCGGACCGGACGAGGCGCGCCAGAGCTGGCCGCA 426
 QY 719 CGGAGAGCTCGCTTCTTCACTCCCGAAGACCGCGAAGGCGCGCGCACG 769
 DB 427 ACCCGAGAGCTGACGAGCGCTTACACCGCCACGCGACGAGCGCGCG 477

RESULT 15
 US-09-565-501A-34
 ; Sequence 34, Application US/09565501A
 ; Patent No. 6607731
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Campos-Neto, Antonio
 ; APPLICANT: Webb, John R.
 ; APPLICANT: Dillion, Davin C.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Bhadia, Ajay
 ; APPLICANT: Coler, Rhea
 ; APPLICANT: Peter Probst
 ; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
 ; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
 ; FILE REFERENCE: 210121.420C6
 ; CURRENT APPLICATION NUMBER: US/09/565,501A
 ; CURRENT FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 112
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 34
 ; LENGTH: 516
 ; TYPE: DNA
 ; ORGANISM: Leishmania major
 ; US-09-565-501A-34

Query Match 4.4%; Score 51.4; DB 4; Length 516;
 Best Local Similarity 47.1%; Pred. No. 0.042;
 Matches 222; Conservative 0; Mismatches 246; Indels 3; Gaps 2;

QY 302 AGCTGAGCGGCGAGTCAACCCAGAGTGGTCAAGCACTGGAGCGCGGCTCGGCCGT 361
 DB 7 AGCAGCGCTTAGACACCGCCACGACGACGCGCGCGAGCTGAGGCAAGGCTGGACGCG 66
 QY 362 GCGCGAGTCCAGCGCGGGGTGTGCGAAGCCGCTGGACGCGA--GACGATCAGCACTGC 420
 DB 67 TGGCGGAGATCGGACGAGCGCGCGCCAGAGCTGGCGCGCAAGCCGAGAGCTGCAGC 126
 QY 421 CACGCGCTGTGTCACACGATGTGCGCGCGCGCGATTCGCGGAAACGATCAGGCTCAAC 480
 DB 127 AGCGCTTAGACACCGCCACGACGACGCGCGCGAGCTGAGGCAAGGCTGGACGCGTGG 186
 QY 481 CCGTGTCTTTCAGACGATGCTGCCCGCGCGCGCGCGAAGAGATGAAGTTCTTAGAGCAC 540
 DB 187 CCGGAGACGCGAGCGAGCGCGCGCGCGCGAGCTGGCGCGAAGCGGAGAGCTGCAGCAGC 246
 QY 541 CCGAGATCGTGGCTTATCAAGGCGCTTCCGCGCGCACTGGCGACCGCTGTATGCTG 600
 DB 247 GCTTAGACACCGCCACGACGAGCGCGCGAGCTGAGGCAAGGTGGCAAGGCTGGCGG 306
 QY 601 CTGATGGAGACCGCTGTAGGTGGGTGAGGCGAT--CGGCTGCGCGCGCGCGGCTG 658
 DB 307 CGAAGCGCGAGAGGTGACAGCGCTTAGACACCGCCACGCGAGCGCGCGCGAGCTGG 366
 QY 659 ACCTGCTGCGCGCGCGCGCGCTGACCGCTGTGAGAGAGCTCCAGAGAGCTGGCCAGCA 718

DB 367 AGGCAAGGGTGGCACGAGCTGGCGCGGACCGCGACGAGCGCGCGCCAGAGCTGGCCGCA 426
 QY 719 CGGAGAGCTGCTTCTTCACTCCCGAAGACCGCGAAGGCGCGCGCACG 769
 DB 427 ACCCGAGAGCTGACGAGCGCTTACACCGCCACGCGAGAGCGCGCGCG 477

Search completed: April 15, 2005, 19:15:51
 Job time : 252.611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 19:01:23 ; Search time 25.1422 seconds
(without alignments)

8204.076 Million cell updates/sec

Title: US-09-855-340B-3

Perfect score: 34
Sequence: 1 ccccgtagcggttcattccatccagtcacccg 34

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5622541 seqs, 303355566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09 NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10 NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11 NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	34	100.0	34	9	US-09-855-340-3
2	34	100.0	247	9	US-09-855-340-6
3	33	97.1	260	9	US-09-855-340-9
4	25	73.5	241	9	US-09-855-340-4
5	25	73.5	243	9	US-09-855-340-5
6	25	73.5	315	9	US-09-855-340-8
7	24	70.6	255	9	US-09-855-340-7
8	22.4	65.9	58073	16	US-10-205-220-1
9	21.4	62.9	623	17	US-10-767-701-4606
10	20.6	60.6	298	18	US-10-424-599-116427
11	20.2	59.4	9025608	15	US-10-156-761-1

12	19.6	57.6	329	17	US-10-424-599-90342
13	19.6	57.6	438	10	US-09-918-995-13891
14	19.6	57.6	556	10	US-09-918-995-15222
15	19.6	57.6	171936	15	US-10-265-071-24
16	19.6	57.6	171936	15	US-10-025-9666-24
17	19.6	57.6	640681	9	US-09-790-988-1
18	19.4	57.1	582	18	US-10-425-115-77062
19	19.4	57.1	595	16	US-10-029-386-7651
20	19.4	57.1	71953	13	US-10-087-192-1057
21	19.4	57.1	352938	18	US-10-322-6966-79
22	19.2	56.5	76	9	US-09-974-300-8405
23	19.2	56.5	76	9	US-09-974-300-8425
24	19.2	56.5	171	9	US-09-974-300-6855
25	19.2	56.5	520	18	US-10-425-115-184239
26	19.2	56.5	839	16	US-10-287-274-7
27	19.2	56.5	987	9	US-09-070-9274-538
28	19.2	56.5	1723	17	US-10-264-237-854
29	19.2	56.5	2239	18	US-10-425-115-161398
30	19.2	56.5	2749	10	US-09-814-353-21697
31	19.2	56.5	41936	17	US-10-052-482-133
32	19.2	56.5	47115	17	US-10-052-482-133
33	19.2	56.5	92219	18	US-10-322-281-805
34	19.2	55.9	582	18	US-10-425-115-153346
35	19.2	55.9	863	18	US-10-425-115-132687
36	19.2	55.9	1581	18	US-10-481-179-1
37	19.2	55.9	1681	10	US-09-892-877-44
38	19.2	55.9	1681	10	US-09-948-783-45
39	19.2	55.9	2364	17	US-10-424-599-18190
40	19.2	55.9	2368	17	US-10-424-599-18190
41	19.2	55.9	24316	13	US-10-087-192-43
42	19.2	55.9	41554	13	US-10-087-192-43
43	18.8	55.3	300	9	US-09-049-695A-25
44	18.8	55.3	300	19	US-10-926-217-25
45	18.8	55.3	596	13	US-10-061-690-2

ALIGNMENTS

RESULT 1
US-09-855-340-3
Sequence 3, Appli
Patent No. US20020076788A1
GENERAL INFORMATION:
APPLICANT: Hosted, Jr., Thomas J.
TITLE OF INVENTION: Isolation of Micromonospora carbonacea var african
TITLE OF INVENTION: PMPI integrase and use of integrating function f
TITLE OF INVENTION: site-specific integration into Micromonospora
TITLE OF INVENTION: halophitica and Micromonospora carbonacea chromos
FILE REFERENCE: IN01164K
CURRENT APPLICATION NUMBER: US/09/855.340
CURRENT FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/204,670
PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 34
TYPE: DNA
ORGANISM: Micromonospora carbonacea
US-09-855-340-3

Query Match 100.0%; Score 34; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ccccgtagcggttcattccatccagtcacccg 34
1 ccccgtagcggttcattccatccagtcacccg 34
RESULT 2

US-09-855-340-6
 ; Sequence 6, Application US/09855340
 ; Patent No. US20020076788A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hosted, Jr., Thomas J.
 ; APPLICANT: Horan, Ann C.
 ; TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
 ; TITLE OF INVENTION: PMPI integrase and use of integrating function for
 ; TITLE OF INVENTION: site-specific integration into Micromonospora
 ; FILE REFERENCE: IN01164K
 ; CURRENT APPLICATION NUMBER: US/09/855,340
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: 60/204,670
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 247
 ; TYPE: DNA
 ; ORGANISM: Micromonospora carbonacea
 US-09-855-340-6

Query Match 100.0%; Score 34; DB 9; Length 247;
 Best Local Similarity 100.0%; Pred. No. 3.8e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCGAGGGTTCAATTCCTCATGTCACCG 34
 DB 92 CCCCCGAGGGTTCAATTCCTCATGTCACCG 125

RESULT 3
 US-09-855-340-9
 ; Sequence 9, Application US/09855340
 ; Patent No. US20020076788A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hosted, Jr., Thomas J.
 ; APPLICANT: Horan, Ann C.
 ; TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
 ; TITLE OF INVENTION: PMPI integrase and use of integrating function for
 ; TITLE OF INVENTION: site-specific integration into Micromonospora
 ; TITLE OF INVENTION: halophytica and Micromonospora carbonacea chromosome
 ; FILE REFERENCE: IN01164K
 ; CURRENT APPLICATION NUMBER: US/09/855,340
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: 60/204,670
 ; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 260
 ; TYPE: DNA
 ; ORGANISM: Micromonospora halophytica
 US-09-855-340-9

Query Match 97.1%; Score 33; DB 9; Length 260;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCGAGGGTTCAATTCCTCATGTCACCG 33
 DB 92 CCCCCGAGGGTTCAATTCCTCATGTCACCG 124

RESULT 4
 US-09-855-340-4
 ; Sequence 4, Application US/09855340
 ; Patent No. US20020076788A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hosted, Jr., Thomas J.
 ; APPLICANT: Horan, Ann C.
 ; TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana

;; TITLE OF INVENTION: PMPI integrase and use of integrating function for
 ;; TITLE OF INVENTION: site-specific integration into Micromonospora
 ;; TITLE OF INVENTION: halophytica and Micromonospora carbonacea chromosome
 ;; FILE REFERENCE: IN01164K
 ;; CURRENT APPLICATION NUMBER: US/09/855,340
 ;; PRIOR FILING DATE: 2001-05-15
 ;; PRIOR APPLICATION NUMBER: 60/204,670
 ;; NUMBER OF SEQ ID NOS: 9
 ;; SOFTWARE: Patentln Ver. 2.1
 ;; SEQ ID NO 4
 ;; LENGTH: 241
 ;; TYPE: DNA
 ;; ORGANISM: Micromonospora carbonacea
 US-09-855-340-4

Query Match 73.5%; Score 25; DB 9; Length 241;
 Best Local Similarity 100.0%; Pred. No. 0.39;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 GGGTTCAATTCCTCATGTCACCG 34
 DB 95 GGGTTCAATTCCTCATGTCACCG 119

RESULT 5
 US-09-855-340-5
 ; Sequence 5, Application US/09855340
 ; Patent No. US20020076788A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hosted, Jr., Thomas J.
 ; APPLICANT: Horan, Ann C.
 ; TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
 ; TITLE OF INVENTION: PMPI integrase and use of integrating function for
 ; TITLE OF INVENTION: site-specific integration into Micromonospora
 ; TITLE OF INVENTION: halophytica and Micromonospora carbonacea chromosome
 ; FILE REFERENCE: IN01164K
 ; CURRENT APPLICATION NUMBER: US/09/855,340
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: 60/204,670
 ; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 243
 ; TYPE: DNA
 ; ORGANISM: Micromonospora carbonacea
 US-09-855-340-5

Query Match 73.5%; Score 25; DB 9; Length 243;
 Best Local Similarity 100.0%; Pred. No. 0.39;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 GGGTTCAATTCCTCATGTCACCG 34
 DB 113 GGGTTCAATTCCTCATGTCACCG 137

RESULT 6
 US-09-855-340-8
 ; Sequence 8, Application US/09855340
 ; Patent No. US20020076788A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hosted, Jr., Thomas J.
 ; APPLICANT: Horan, Ann C.
 ; TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
 ; TITLE OF INVENTION: PMPI integrase and use of integrating function for
 ; TITLE OF INVENTION: site-specific integration into Micromonospora
 ; TITLE OF INVENTION: halophytica and Micromonospora carbonacea chromosome
 ; FILE REFERENCE: IN01164K
 ; CURRENT APPLICATION NUMBER: US/09/855,340
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: 60/204,670

PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 315
TYPE: DNA
ORGANISM: Micromonospora halophytica
US-09-855-340-8

Query Match 73.5%; Score 25; DB 9; Length 315;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 GGGTCAATTCCTCAGTCAACCC 34
Db 96 GGGTCAATTCCTCAGTCAACCC 120

RESULT 7
US-09-855-340-7
Sequence 7, Application US/09855340
Patent No. US20020076788A1
GENERAL INFORMATION:
APPLICANT: Hosted, Jr., Thomas J.
TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
TITLE OF INVENTION: PMPI integrase and use of integrating function for
TITLE OF INVENTION: site-specific integration into Micromonospora
TITLE OF INVENTION: halophytica and Micromonospora carbonacea chromosome
FILE REFERENCE: IN01164K
CURRENT APPLICATION NUMBER: US/09/855,340
CURRENT FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/204,670
PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 255
TYPE: DNA
ORGANISM: Micromonospora halophytica
US-09-855-340-7

Query Match 70.6%; Score 24; DB 9; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 GGGTCAATTCCTCAGTCAACCC 33
Db 96 GGGTCAATTCCTCAGTCAACCC 119

RESULT 8
US-10-205-220-1

Sequence 1, Application US/10205220
Publication No. US20030170663A1
GENERAL INFORMATION:
APPLICANT: Fraser et al.
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
TITLE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB193P1D1
CURRENT APPLICATION NUMBER: US/10/205,220
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: US 08/545,528
PRIOR FILING DATE: 1995-10-19
PRIOR APPLICATION NUMBER: US 08/488,018
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/473,545
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 580073
TYPE: DNA

ORGANISM: Mycoplasma genitalium
US-10-205-220-1

Query Match 65.9%; Score 22.4; DB 16; Length 580073;
Best Local Similarity 95.8%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 10 GGGTCAATTCCTCAGTCAACCC 33
Db 445128 GGGTCAATTCCTCAGTCAACCC 445151

RESULT 9
US-10-767-701-4606/c
Sequence 4606, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associa
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 4606
LENGTH: 623
TYPE: DNA
ORGANISM: Sorghum bicolor
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS839651
US-10-767-701-4606

Query Match 62.9%; Score 21.4; DB 18; Length 623;
Best Local Similarity 80.6%; Pred. No. 18;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 3 CCGGTACGGTCAATTCCTCAGTCAACCC 33
Db 382 CCGGTACGGTCAATTCCTCAGTCAACCC 352

RESULT 10
US-10-424-599-116427
Sequence 116427, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Ass
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 116427
LENGTH: 298
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(298)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_76144C.1
US-10-424-599-116427

Query Match 60.6%; Score 20.6; DB 17; Length 298;
Best Local Similarity 85.2%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Mon Apr 18 09:47:32 2005

us-09-855-340b-3.rnpb

RESULT 15
US-10-265-071-24

; Sequence 24, Application US/10265071
; Publication No. US20030147875A1

; GENERAL INFORMATION:

; APPLICANT: Rosen, Steven D.

; APPLICANT: Hemmerich, Stefan

; APPLICANT: Tomita, Megumi

; TITLE OF INVENTION: Sulfatases and methods of use thereof

; FILE REFERENCE: UCAL-230CIP

; CURRENT APPLICATION NUMBER: US/10/265,071

; CURRENT FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,577

; PRIOR FILING DATE: 2000-12-27

; PRIOR APPLICATION NUMBER: 60/267,831

; PRIOR FILING DATE: 2001-09-02

; PRIOR APPLICATION NUMBER: 10/025,966

; PRIOR FILING DATE: 2001-12-21

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 24

; LENGTH: 171936

; TYPE: DNA

; ORGANISM: Mus musculus

; US-10-265-071-24

Query Match 57.6%; Score 19.6; DB 15; Length 171936;
Best Local Similarity 84.6%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 9 CGGTTCAATTCGCATGTCACCG 34

DB 20349 CGAGTCAATTCGCATGTCACCG 20374

Search completed: April 15, 2005, 22:51:32
Job time: 40.1422 secs

THIS PAGE BLANK (05/10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2005, 17:52:06 ; Search time 7.19829 Seconds
(without alignments)
7728.696 Million cell updates/sec

Title: US-09-855-340B-3

Perfect score: 34
Sequence: 1 ccccggtacgggttcattcccatcagtcaccgcg 34

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query Match Length	DB ID	Description
1	22.4	65.9 580073	4 US-08-545-528D-1	Sequence 1, Appli
2	20.8	61.2 1781	4 US-09-902-540-180	Sequence 180, App
3	19.6	57.6 640681	4 US-09-790-988-1	Sequence 1, Appli
4	19.2	56.5 306	4 US-09-134-000C-892	Sequence 892, App
5	19.2	56.5 839	4 US-09-711-164-7	Sequence 7, Appli
6	19.2	56.5 1203	4 US-09-902-540-9122	Sequence 9122, Ap
7	19.2	56.5 10276	4 US-09-902-540-975	Sequence 975, App
8	18.8	55.3 596	4 US-08-815-175-2	Sequence 2, Appli
9	18.8	55.3 3037	4 US-09-573-080A-42	Sequence 256, App
10	18.8	55.3 6305	3 US-09-221-017B-256	Sequence 6, Appli
11	18.8	55.3 9511	1 US-07-925-695-6	Sequence 7, Appli
12	18.8	55.3 9511	1 US-07-925-695-7	Sequence 28, Appli
13	18.6	54.7 1538	4 US-09-644-460-28	Sequence 10, Appli
14	18.6	54.7 2463	1 US-08-307-444A-10	Sequence 1, Appli
15	18.6	54.7 2463	1 US-08-587-389-10	Sequence 10, Appli
16	18.6	54.7 3466	4 US-09-880-484D-1	Sequence 1, Appli
17	18.6	54.7 3466	4 US-10-438-648-1	Sequence 1326, Ap
18	18.6	54.7 3653	4 US-09-023-655-1326	Sequence 17, Appli
19	18.6	54.7 35100	2 US-08-770-379-17	Sequence 17, Appli
20	18.6	54.7 35100	3 US-08-757-669A-17	Sequence 17, Appli
21	18.6	54.7 35100	3 US-09-230-371A-17	Sequence 17, Appli
22	18.6	54.7 4403765	3 US-09-103-840A-2	Sequence 2, Appli
23	18.6	54.7 4411529	3 US-09-103-840A-1	Sequence 32, Appli
24	18.4	54.1 1894	3 US-09-329-350-32	Sequence 32, Appli
25	18.4	54.1 1894	4 US-08-841-636A-32	Sequence 32, Appli
26	18.2	53.5 601	4 US-09-949-016-56512	Sequence 56512, A
27	18.2	53.5 1281	4 US-09-489-039A-6674	Sequence 6674, Ap

28	18.2	53.5	1356	4 US-09-489-039A-359	Sequence 359,
29	18.2	53.5	2751	1 US-08-153-848-23	Sequence 23,
30	18.2	53.5	2751	3 US-09-299-843A-23	Sequence 23,
31	18.2	53.5	2751	3 US-09-088-337B-23	Sequence 23,
32	18.2	53.5	2751	5 PCT-US93-11153-23	Sequence 23,
33	18.2	53.5	5039	4 US-09-386-816C-1	Sequence 1, A
34	18.2	53.5	5039	4 US-10-320-176-1	Sequence 1, A
35	18.2	53.5	17657	4 US-09-949-016-16135	Sequence 16133
36	18.2	53.5	143644	4 US-09-949-016-15238	Sequence 1523
37	18.2	53.5	187848	4 US-09-949-016-12111	Sequence 1211
38	18.2	53.5	636591	4 US-09-949-016-11808	Sequence 1180
39	18.2	53.5	636591	4 US-09-949-016-13368	Sequence 1338
40	18	52.9	472	4 US-09-548-938A-9	Sequence 9, A
41	18	52.9	4451	4 US-09-548-938A-1	Sequence 1, A
42	18	52.9	68924	4 US-09-949-016-15367	Sequence 1536
43	18	52.9	79835	4 US-09-949-016-12456	Sequence 1245
44	18	52.9	79835	4 US-09-949-016-16121	Sequence 1612
45	18	52.9	119801	4 US-09-949-016-13453	Sequence 1345

ALIGNMENTS

RESULT 1
US-08-545-528D-1
Sequence 1, Application US/08545528D
Patent No. 6537773
GENERAL INFORMATION:
APPLICANT: Frazer et al.
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium G
Patent No. 6537773
TITLE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB193p1
CURRENT FILING DATE: 1995-10-19
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/488, 018
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/473, 545
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 580073
TYPE: DNA
ORGANISM: Mycoplasma genitalium
US-08-545-528D-1
Query Match 65.9%; Score 22.4; DB 4; Length 580073;
Best Local Similarity 95.8%; Pred. No. 7.3;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps
DB 445128 GGGTCAATCCCATCAGTCCGCC 445151
QY 10 GGGTCAATCCCATCAGTCCGCC 33
DB 445128 GGGTCAATCCCATCAGTCCGCC 445151
RESULT 2
US-09-902-540-180
Sequence 180, Application US/0902540
Patent No. 683447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses There
FILE REFERENCE: 38-10(15849)B
CURRENT FILING DATE: 2001-07-10
PRIOR FILING DATE: 2000-07-10
PRIOR APPLICATION NUMBER: 60/217, 883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 180

LENGTH: 1781
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-180

Query Match 61.2%; Score 20.8; DB 4; Length 1781;
Best Local Similarity 91.7%; Pred. No. 8.1;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 GGGTTCATTCCCATCAGTCACCC 33
DB 1396 GGGTTCATTCCCATCAGTCACCC 1419

RESULT 3
US-09-988-1/c
Sequence 1, Application US/09790988
Patent No. 6632935

GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: MATANABE, HIDEKI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 57.6%; Score 19.6; DB 4; Length 640681;
Best Local Similarity 84.6%; Pred. No. 1.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 ACCGGTTCATTCCCATCAGTCACCC 33
DB 625858 ATGGGTTCAAGTCCATAGCCACCC 625833

RESULT 4
US-09-134-000C-892/c
Sequence 892, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucet-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 892
LENGTH: 306
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-892

Query Match 56.5%; Score 19.2; DB 4; Length 306;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10 GGGTTCATTCCCATCAGTCACCC 33
DB 10 GGGTTCATTCCCATCAGTCACCC 33

DB 251 GGGTTCATTCCCATCAGTCACCC 228

RESULT 5
US-09-711-164-7

Sequence 7, Application US/09711164
Patent No. 6589738
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allyn
APPLICANT: Olsen, Kari
APPLICANT: Zyskind, Judith
TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THEREOF
FILE REFERENCE: ELITRA.008A
CURRENT APPLICATION NUMBER: US/09/711,164
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/164415
PRIOR FILING DATE: 1999-11-9
NUMBER OF SEQ ID NOS: 469
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 839
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(839)
OTHER INFORMATION: n = A,T,C or G
US-09-711-164-7

Query Match 56.5%; Score 19.2; DB 4; Length 839;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CCCGTACGGTTCATTCCCATCAGTCACCG 34
DB 640 CNGGTCAGGTTCAATTACCATCGTAATCG 672

RESULT 6
US-09-902-540-9122/c
Sequence 9122, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 9122
LENGTH: 1203
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-9122

Query Match 56.5%; Score 19.2; DB 4; Length 1203;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 CCGTACGGTTCATTCCCATCAGTCACCG 34
DB 1098 CCGTACGGTTCATTCCCATCAGTCACCG 1067

RESULT 7
US-09-902-540-975
Sequence 975, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Wiegand, Roger C.
 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 FILE REFERENCE: 38-10(15849)B
 CURRENT APPLICATION NUMBER: US/09/902,540
 CURRENT FILING DATE: 2001-07-10
 PRIOR APPLICATION NUMBER: 60/217,883
 PRIOR FILING DATE: 2000-07-10
 NUMBER OF SEQ ID NOS: 16825
 SEQ ID NO 975
 LENGTH: 10276
 TYPE: DNA
 ORGANISM: Myxococcus xanthus
 US-09-902-540-975

Query Match 56.5%; Score 19.2; DB 4; Length 10276;
 Best Local Similarity 75.0%; Pred. No. 74;
 Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 CCGGTACGGGTTCAATCCCATCAGTACCCG 34
 DB 1945 CCGGTACGGGTTGTAAGCGCCAGCGGTACCCG 1976

RESULT 8
 US-08-815-175-2
 Sequence 2, Application US/08815175
 Patent No. 5856139
 GENERAL INFORMATION:
 APPLICANT: Lal, Preeti
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Goli, Surya K.
 TITLE OF INVENTION: NOVEL PROLINE-RICH ACIDIC PROTEIN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: US
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/815,175
 FILING DATE: Filed Herewith
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0225 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 596 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: PANTCUT02
 CLONE: 2235738
 US-08-815-175-2

Query Match 55.3%; Score 18.8; DB 2; Length 596;
 Best Local Similarity 69.7%; Pred. No. 51;
 Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CCGGTACGGGTTCAATCCCATCAGTACCC 33
 DB 497 CCGGTACGGGTTGTAAGCGCCAGCGGTACCC 529

RESULT 9
 US-09-573-080A-42/c
 Sequence 42, Application US/09573080A
 Patent No. 6828097
 GENERAL INFORMATION:
 APPLICANT: JOAN, KNOLL
 APPLICANT: ROGAN, PETER
 TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHODS
 FILE REFERENCE: 30307
 CURRENT APPLICATION NUMBER: US/09/573,080A
 CURRENT FILING DATE: 2000-05-16
 NUMBER OF SEQ ID NOS: 479
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 42
 LENGTH: 3037
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: repeat region
 LOCATION: (1)..(3037)
 OTHER INFORMATION: hev168
 NAME/KEY: misc feature
 OTHER INFORMATION: n is a, c, g or t
 PUBLICATION INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS: Jurka, J; Walichewicz, J; Miosavljovic, A
 TITLE: Prototypic sequences for human repetitive DNA
 JOURNAL: Journal of Molecular Evolution
 VOLUME: 35
 ISSUE: 4
 PAGES: 286-291
 DATE: 1992-10-
 DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
 DATABASE ENTRY DATE: 1996-01-26
 DATABASE ENTRY DATE: 1996-01-26
 US-09-573-080A-42

Query Match 55.3%; Score 18.8; DB 4; Length 3037;
 Best Local Similarity 76.7%; Pred. No. 81;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCGGTACGGGTTCAATCCCATCAGTCA 30
 DB 1713 CCGGTACGGGTTGTAAGTCAATCTGTCA 1684

RESULT 10
 US-09-221-017B-256/c
 Sequence 256, Application US/09221017B
 Patent No. 6444799
 GENERAL INFORMATION:
 APPLICANT: Ross, Bruce C.
 TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
 NUMBER OF SEQUENCES: 1120
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 PAGE MILL ROAD
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 256:
SEQUENCE CHARACTERISTICS:
LENGTH: 6305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...6305
US-09-221-017B-256

Query Match 55.3%; Score 18.8; DB 3; Length 6305;
Best Local Similarity 76.7%; Pred. No. 99;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

4 CGGTACGGGTTCATTCCTCAGTCACTCACC 33
2051 CGGTACATATTCGTTTACCATCAGTCCGCC 2022

RESULT 11
US-07-925-695-6
Sequence 6, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
APPLICANT: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, Degrandi, Weillacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/925,695
FILING DATE: 19920807
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Weillacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/87-48009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
TELEX: WUI 64470
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9511 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-925-695-6

Query Match 55.3%; Score 18.8; DB 1; Length 9511;
Best Local Similarity 66.7%; Pred. No. 1,1e+02;
Matches 20; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

2 CCGGTACGGGTTCATTCCTCAGTCACTCACC 31
4407 CCGGTACGGGTTCATTCCTCAGTCACTCACC 4436

RESULT 12
US-07-925-695-7
Sequence 7, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
APPLICANT: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, Degrandi, Weillacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/925,695
FILING DATE: 19920807
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Weillacher, Robert G.
REGISTRATION NUMBER: 20,531

REFERENCE/DOCKET NUMBER: 06/87-48009

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 659-2811

TELEFAX: (202) 659-1462

TELEX: WUI 64470

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 9511 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

US-07-925-695-7

Query Match

Best Local Similarity 55.3%; Score 18.8; DB 1; Length 9511;

Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 4407 CCCGTAACGGGTTCATTCCTCCATCAGTCAC 31

4407 CCCGTAACGGGTTCATTCCTCCATCAGTCAC 4436

RESULT 13

US-09-644-460-28/C

Sequence 28, Application US/09644460

Patent No. 6657053

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.

TITLE OF INVENTION: Reciprocal Subtraction Differential

FILE REFERENCE: 34587-C-PCT-USA

CURRENT FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: US/09/644,460

PRIOR FILING DATE: 1999-02-26

PRIOR APPLICATION NUMBER: PCT/US99/04323

PRIOR FILING DATE: 1998-11-23

PRIOR APPLICATION NUMBER: US 09/185,115

PRIOR FILING DATE: 1998-11-03

PRIOR APPLICATION NUMBER: US 09/032,684

PRIOR FILING DATE: 1998-02-27

NUMBER OF SEQ ID NOS: 42

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 28

LENGTH: 1538

TYPE: DNA

ORGANISM: homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: 652, 1523

OTHER INFORMATION: c, t, a or g

NAME/KEY: misc_feature

LOCATION: (1)...(1538)

OTHER INFORMATION: n = A,T,C or G

US-09-644-460-28

Query Match

Best Local Similarity 54.7%; Score 18.6; DB 4; Length 1538;

Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1240 CCCGTAACGGGTTCATTCCTCCATCAGTCACCC 33

1240 CCCGTAACGGGTTCATTCCTCCATCAGTCACCC 1208

RESULT 14

US-08-307-444A-10

Sequence 10, Application US/08307444A

Patent No. 5516659

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI

APPLICANT: MORISHITA, HIDEAKI

APPLICANT: UEMURA, AKIO

APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/06/307,444A

FILING DATE: 19-SEP-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/835,436

FILING DATE: 26-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: OLIF, JAMES A.

REGISTRATION NUMBER: 27,075

REFERENCE/DOCKET NUMBER: JAO 27706

TELEPHONE: (703) 836-6400

TELEFAX: (703) 836-2787

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 2463 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-307-444A-10

Query Match

Best Local Similarity 54.7%; Score 18.6; DB 1; Length 2463;

Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 2338 CCCGTAACGGGTTCATTCCTCCATCAGTCACCC 33

2338 CCCGTAACGGGTTCATTCCTCCATCAGTCACCC 2370

RESULT 15

US-08-587-389-10

Sequence 10, Application US/08587389

Patent No. 5695964

GENERAL INFORMATION:

APPLICANT: NII, ATSUSHI

APPLICANT: MORISHITA, HIDEAKI

APPLICANT: UEMURA, AKIO

APPLICANT: MOCHIDA, EI

TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT

NUMBER OF SEQUENCES: 27

PRODUCTION METHOD: AND THERAPEUTIC AGENT (AS AM

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIF & BERRIDGE

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/587,389
 FILING DATE: 17-JAN-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/307,444
 FILING DATE: 19-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: OLIEF, JAMES A.
 REGISTRATION NUMBER: 27,075
 REFERENCE/DOCKET NUMBER: JAO 27706
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6400
 TELEFAX: (703) 836-2787
 TELEX: 90-1799 PTO ALEX
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2463 bases
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-587-389-10

Query Match 54.7% Score 18.6; DB 1; Length 2463;
 Best Local Similarity 72.7%; Pred. No. 95;
 Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Oy 1 CCCCGTACGGGTTCATTCCTCATCTCAGTCAACC 33
 Db 2338 CCCAGAACTGGGCACTCACCCTGAGTCAACC 2370

Search completed: April 15, 2005, 19:16:05
 Job time: 15.1983 secs